



# BIREDIAL-ISTEC

3-7 de octubre 2022 COSTA RICA

Actas de la XI Conferencia Internacional de  
Bibliotecas y Repositorios Digitales



Red Mexicana  
de Repositorios  
Institucionales



# XI Conferencia Internacional de Bibliotecas y Repositorios Digitales

Modalidad virtual

3 al 7 de octubre de 2022



Sede

Costa Rica

Actas de la XI Conferencia Internacional de Bibliotecas y Repositorios Digitales /  
Rebeca Marín del Campo... [et al.] ; coordinación general de Alexa Ramírez Vega. -  
1a ed. - La Plata : Universidad Nacional de La Plata. SEDICI, 2023.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-950-34-2219-9

1. Acceso a la Información. I. Marín del Campo, Rebeca. II. Ramírez Vega, Alexa, coord.

CDD 020.7

Coordinación general de la edición: Ing. ALEXA RAMÍREZ VEGA

Edición, maquetación y corrección de estilo: Esp. ANALÍA V. PINTO

Diseño de tapa: DCV LUCAS E. FOLEGOTTO

Esta obra se edita bajo una licencia [CC-BY](https://creativecommons.org/licenses/by/4.0/).



## COMITÉ CIENTÍFICO

CATERINA GROPOSO, Universidade Federal do Rio Grande do Sul (Brasil)

MALGORZATA LISOWSKA, Universidad del Rosario (Colombia)

MARISA R. DE GIUSTI, Universidad Nacional de La Plata (Argentina)

MARLEM URIBE MARENCO, Universidad del Norte (Colombia)

MEILYN GARRO, Universidad de Costa Rica (Costa Rica)

ROSALINA VÁZQUEZ, Red Mexicana de Repositorios Institucionales (México)

## SEDE 2022

El Consejo Nacional de Rectores fue creado mediante Convenio de Coordinación de la Educación Superior Universitaria Estatal en Costa Rica, suscrito por las Instituciones de Educación Superior Universitaria Estatal, el 4 de diciembre de 1974. En él se regulan aspectos de coordinación para el ejercicio conjunto de la autonomía universitaria en diversos ámbitos. Las universidades miembros de CONARE son: la Universidad de Costa Rica, el Instituto Tecnológico de Costa Rica, la Universidad Nacional, la Universidad Estatal a Distancia y la Universidad Técnica Nacional.

Desde hace más de diez años, el Consejo impulsa iniciativas relacionadas con el acceso abierto, el mejoramiento de la calidad de las publicaciones académicas y la promoción de los repositorios institucionales.

Debido a posibles cambios en las medidas sanitarias relacionadas con la pandemia COVID-19, se tomó la decisión de realizar un evento virtual y así dar oportunidad de presentar sus propuestas a la mayor cantidad de participantes.

Bienvenidos a la conferencia de este año.

## COMITÉ ORGANIZADOR LOCAL

**ALEXA RAMÍREZ VEGA, Instituto Tecnológico de Costa Rica**

**ANDREA MARÍN CAMPOS, Universidad de Costa Rica**

**ANDREA MÉNDEZ SOLANO, Universidad Técnica Nacional**

**ANDREA MORA CAMPOS, Universidad Nacional**

**CAROLINA SEAS CARVAJAL, Universidad Estatal a Distancia**

**EDWARD ARAYA HIDALGO, Consejo Nacional de Rectores**

**FABIOLA CAMPOS JARA, Universidad Nacional**

**JASON ENRIQUEZ FALLAS, Instituto Tecnológico de Costa Rica**

**JORGE POLANCO CORTÉS, Universidad de Costa Rica**

**MEILYN GARRO ACÓN, Universidad de Costa Rica**

**SHARLÍN SÁNCHEZ ESPINOZA, Consejo Nacional de Rectores**

**STEVEN SEGURA JIMÉNEZ, Universidad Estatal a Distancia**

# Índice

PRESENTACIÓN	9
ÉNFASIS TEMÁTICO 2022	10
CONFERENCIAS MAGISTRALES	11
¿Qué es la gestión de datos y cómo la relaciono con mi proyecto de investigación?	12
El presente y el futuro del Open Journal Systems (OJS)	15
Evaluación responsable de la investigación en la transición hacia la ciencia abierta en América Latina y el Caribe. CLACSO-FOLEC	18
Gestión de la calidad y la preservación en repositorios institucionales	22
Mesa de debate: Experiencias en el desarrollo de Repositorios Nacionales y Regionales de Patrimonio Cultural de América Latina	29
Gestión de Recursos Educativos Abiertos y Prácticas Educativas Abiertas Recomendaciones UNESCO 2019	40
¿Qué es la Research Data Alliance?	44
Consortia Papyrus – Repositorio de Datos de Investigación en Colombia, una iniciativa de trabajo colaborativo	50
Políticas institucionales de gestión de datos: buenas prácticas, retos y casos en Brasil, Colombia y Argentina	54
Mesa de debate: Evaluación abierta en revistas	59
PONENCIAS	64
La innovación en la enseñanza y el aprendizaje de las artes visuales y plásticas en estudiantes de la EPEA-UJED	66
Produção científica sobre os princípios FAIR a partir do Directory of Open Access Journal (DOAJ)	83

Políticas nacionais de apoio à ciência aberta: análise contrastiva das iniciativas desenvolvidas nos Estados Unidos, Portugal e Brasil _____	98
Transformación digital de un archivo histórico mediante el uso de un software de código abierto: el caso del Archivo Histórico de la Real Caja de Zacatecas ____	129
Comunicación de la ciencia, la tecnología y la innovación: elementos para una conceptualización democratizadora _____	144
De la gestión de archivos sonoros: teorías y conceptos _____	163
A importância do povoamento da produção acadêmica de TCCs/TCRs no Arca – Repositório Institucional da Fiocruz _____	180
Repensar los Recursos Educativos Abiertos desde los paradigmas de la educación _____	189
MOOC UTN: una propuesta abierta, masiva y en línea _____	204
Gestión de los datos para la investigación en los repositorios del área de ciencias de la tierra de la UNAM _____	214
Planejamento e incubação de repositórios de dados em ICTs brasileiras _____	233
HERA - Herramienta para Enriquecimiento de Recursos Académicos _____	251
SUM-OJS: Scripts para actualizaciones y migraciones de OJS _____	268
Es necesario un identificador persistente además del DOI (Digital Object Identifier) para el sur global _____	292
Evaluación de un repositorio institucional a través de NDSA Levels: Caso CIC Digital _____	304
Experiencias en la migración de un repositorio institucional a DSpace 7 ____	322
WebID como base para el desarrollo de una marca personal en repositorios institucionales _____	341
Conectando fortalezas y capacidades institucionales: automatización de la integración VIVO-Pure-DSpace _____	358
PÓSTERS _____	370
Caracterización de las revistas de acceso abierto en la Universidad Nacional, Costa Rica, 2022 _____	371

RedELCA: una propuesta para visibilizar la producción literaria y cultural en Centroamérica y el Caribe _____	373
Comunicación pública de la ciencia en la Universidad de Cienfuegos: políticas y acciones _____	375
Gestión de la investigación y visibilidad de la producción científica y normativa en el Instituto de Evaluación de Tecnologías en Salud e Investigación IETSI, EsSalud. Lima- Perú, 2021-2022 _____	377
Ciencia Abierta en instituciones de educación superior estatal: estrategias y experiencias de la Subcomisión de Conocimiento Abierto del Consejo Nacional de Rectores (CONARE) de Costa Rica _____	379
Acceso al patrimonio cultural digital el caso del Repositorio de Patrimonio Cultural Centroamericano _____	381
Planificación e implementación del Repositorio de Datos Académicos de la Universidad Nacional de Rosario (RDA-UNR): una metodología colaborativa e interdisciplinaria _____	382
Retos de la evaluación de la investigación: Caso de las universidades públicas costarricenses _____	385

# Presentación

Desde su fundación en 2011, el objetivo general de la Conferencia BIREDIAL-ISTEC ha sido promover la iniciativa de Acceso Abierto, la visibilidad institucional latinoamericana y, a partir de 2021, se adquiere un compromiso con la promoción de los demás componentes de Ciencia Abierta.

El público al que está dirigida la Conferencia está compuesto por:

- Comunidades científicas, observatorios, bibliotecas, centros de datos abiertos y otros
- Profesionales interesados en cuestiones relacionadas con la gestión, promoción y visibilidad del conocimiento
- Autoridades académicas
- Investigadores, editores y directores de publicaciones científicas, alumnos de varias especialidades
- Instituciones de enseñanza superior y centros de investigación
- Personal del gobierno y de la administración pública interesada en compartir preocupaciones relacionadas con la transparencia de los datos públicos

La **XI Conferencia Internacional BIREDIAL-ISTEC** se realizó del 3 al 7 de octubre de 2022 en modalidad virtual desde Costa Rica, a cargo del Consejo Nacional de Rectores (CONARE), conformado por la Universidad de Costa Rica (UCR), el Instituto Tecnológico de Costa Rica (TEC), la Universidad Nacional (UNA), la Universidad Estatal a Distancia (UNED) y la Universidad Técnica Nacional (UTN).

## Énfasis temático 2022

El desarrollo de métricas alternativas y políticas en la búsqueda de diversificar los indicadores de evaluación científica.

BIREDIAL-ISTEC 2022 presentará el debate sobre cómo los criterios y políticas impactan la forma en que la ciencia es concebida actualmente, y un espacio para dialogar y analizar el desarrollo de métricas alternativas y políticas que permitan diversificar los indicadores utilizados para calificar a las personas investigadoras, en busca de una evaluación más inclusiva y justa con las comunidades académicas de todos los campos disciplinares.

Los criterios de evaluación a personas investigadoras afectan directamente la manera en cómo la ciencia se distribuye, y la forma en que se asignan los fondos para realizar investigación. A pesar de que la comunidad académica de diferentes partes del mundo ha validado los beneficios de la ciencia abierta y los riesgos de las métricas tradicionales, esta última sigue siendo el modelo más utilizado para evaluar y asignar fondos a nivel mundial.



---

Conferencias  
magistrales

---

# ¿Qué es la gestión de datos y cómo la relaciono con mi proyecto de investigación?

MESA DE DEBATE

28 DE SEPTIEMBRE DE 2022

## **REBECA MARÍN DEL CAMPO**

Universidad Carlos III de Madrid (España)

## **EVA ORTIZ**

Universidad Carlos III de Madrid (España)

### **RESUMEN EXTENDIDO**

Las expositoras hacen un recorrido en torno a los datos de investigación, los tipos existentes y sus conjuntos. Para esto proponen ejemplos como es el Proyecto Genoma Humano, en el cual las bases de datos se hallan a disposición de los científicos en el mundo, sin restricciones en cuanto a su uso o redistribución.

Basándose en la importancia de los datos en abierto y la utilidad de compartirlos, refieren la recomendación de la UNESCO (2021), lo que implica un cambio de paradigma en la gestión de los datos de investigación. Y mencionan que los datos en abierto verifican los resultados de investigación, permiten su reutilización; aumentan la credibilidad de la ciencia, pues hace patente la inversión pública en investigación; mejoran de la transparencia; aumentan la visibilidad y el impacto de la investigación; favorecen la generación de nuevos contactos entre colegas, impulsando la colaboración y finalmente contribuyen en la preservación.

Otro tema al que hacen referencia es al acrónimo FAIR, que en español se traduce como: “Encontrable”, “Accesible”, “Interoperable”, “Reutilizable”, establecidos por la iniciativa Force 11 para determinar cómo deben ser los datos al compartirse.

Posteriormente, se hace mención al tema de las políticas, las cuales resultan fundamentales porque crean un marco de actuación y además potencian el desarrollo y la implantación de la gestión de los datos. Además, las expositoras mencionan algunas iniciativas en el ámbito de la Unión Europea, como Horizonte Europa, la European Open Science Cloud, la Ley de España 17/2022 de la Ciencia, la Tecnología y la Innovación, que en su artículo 37 que hace referencia a la Ciencia Abierta. Y a nivel latinoamericano, las Políticas para la Ciencia Abierta y los Datos Científicos en América Latina y legislaciones nacionales de Perú, Argentina, México y proyectos en Chile y Brasil.

Tal información se complementa con el ciclo de vida de los datos de investigación, que corresponde con la gestión de datos de investigación (GDI) y, a su vez, implica una serie de aspectos tales como la propiedad y los tipos de datos, el almacenamiento, los formatos, las normas disciplinarias de metadatos, la estructura y convenciones, y la documentación. También se evidencian aspectos éticos y datos personales, para el uso adecuado de los datos «sensibles», y las buenas prácticas como la anonimización, que supone el procesamiento de los datos personales de forma que se prevenga de manera irreversible la identificación. Puede ser por razones éticas, como proteger la identidad.

Considerando los aspectos anteriores se puede proceder a la preparación de un Plan de Gestión de Datos (DMP/PGD), que es un documento formal que debe presentarse al inicio de la investigación y modificarse según se avanza con el proyecto, por lo tanto, es un documento vivo. Existen varias plataformas de apoyo para cumplimentarlo, todas basadas en un cuestionario donde explicar la gestión durante el ciclo de vida de los datos, desde su creación hasta la finalización del proyecto de investigación.

Algunos ejemplos de plantillas normalizadas para diferentes financiados (son de uso libre):

- Horizonte Europa: [Plantilla para DMP](#)
- PARTHENOS: [Plantilla para datasets de arqueología](#)
- Digital Curation Centre: [Checklist for a Data Management Plan](#)
- [Madroño PGD Online](#)
- [Argos](#)

Existen también buenas prácticas para el tratamiento de ficheros a depositar, como son: nombres, formatos, protección de datos personales, etc. para que estos sean significativos/descriptivos y consistentes. La recomendación es aplicarlas tanto para nombres, como para los distintos formatos.

El siguiente aspecto es el de las licencias de uso, que corresponden a la forma de comunicar qué permisos disponen los potenciales usuarios, las cuales están determinadas por las personas creadoras, productoras o distribuidoras de los datos. Para identificar cuáles son los tipos se puede acudir a Creative Commons, Open Data Commons, y para licenciar software a MIT License, Apache License, y GNU General Public License.

En cuanto al tema de los repositorios, se considera para la difusión de datos finales que sean «tan abiertos como sea posible, tan cerrados como sea necesario» y visualizar la importancia de los repositorios en la preservación.

Finalmente, se presentan ejemplos concretos de la experiencia en España, como el Consorcio Medroño, en el cual se han elaborado guías de cómo depositar los datos para el personal investigador de las universidades que forman parte del Consorcio, o el Grupo Técnico Iberoamericano de Dataverse conformado por universidades y consorcios. Y otro ejemplo en la Universidad Carlos III que también pone a disposición una serie de indicaciones para el depósito de datos en e-ciencia de datos.

[Ver conferencia en YouTube](#)

# El presente y el futuro del Open Journal Systems (OJS)

CONFERENCIA MAGISTRAL

28 DE SEPTIEMBRE 2022

## JUAN PABLO ALPERÍN

Universidad de Simon Fraser (Canadá)

### RESUMEN EXTENDIDO

El objetivo de esta conferencia es evidenciar la importancia del uso de Open Journal Systems (OJS), para así poder evidenciar el volumen y el impacto que está teniendo PKP y OJS en el mundo, en términos de ayudar y apoyar a la comunicación científica global, pero más que poder resaltar el volumen de la comunicación científica que existe en el mundo fuera de los circuitos *mainstream* y de las bases de datos comerciales, que en América Latina son tan utilizadas para el momento de hacer evaluación, pero que lamentablemente ignoran que un alto porcentaje de la ciencia existe en el mundo en su intento por hacerse notar.

Conforme a algunas estadísticas que se han realizado hasta el momento, hay un gran porcentaje ahora, más de la mitad que están publicando en Asia y en particular en Indonesia, lo cual representa un cambio dramático en los últimos años, en los que por mucho tiempo América Latina ocupaba este lugar.

Sin embargo, existen más de cinco mil revistas en América Latina con OJS, lo que sugiere una importante influencia en lo que hace PKP y es una comunidad más consolidada y con la que tenemos mucho más contacto que con la de Asia.

En términos de diversidad, es importante resaltar que gran cantidad de revistas se están publicando en inglés, un gran porcentaje en indonesio, pero el español y el portugués son los otros idiomas más utilizados, que suman más del veinte por ciento de las revistas que se están utilizando en el sistema. Lo cual apunta al tema del multilingüismo, donde se ha identificado que un poco más de la mitad de las revistas que utilizan OJS publican en dos o tres idiomas.

No es cosa menor comprender que el rol que está teniendo OJS y PKP en la región es no solamente el de apoyar a revistas que están tratando de acoplarse al *mainstream* de la ciencia o que están tratando de cumplir con los «estándares globales». Lo que interesa resaltar es que hay una diversidad de actividades en términos de los países que están logrando publicar en idiomas como el náhuatl, quechua, etc., que responden a historias locales de países, es decir, las revistas que se están publicando o utilizando OJS no son solamente aquellas que están tratando de vincularse con una ciencia *mainstream* o global, sino que en muchos casos son revistas que están tratando temas que no encontrarían una voz o no encontrarían otra salida más que a través de revistas con apoyos locales.

Así mismo, hay una gran cantidad de revistas pequeñas, distribuidas con apoyo institucional. Son revistas que generalmente son independientes, y no son parte de una editorial comercial o guiadas por académicos, están distribuidas en todo el mundo y a través de muchas instituciones. Cada instalación tiene en promedio dos revistas y media, es decir, no se trata de instalaciones enormes, y la mayoría tienen algún apoyo institucional, generalmente instituciones académicas, universidades especialmente, pero algunas también de asociaciones.

Cuando las revistas están disponibles al público tienen un valor muy importante, y esto en algunos circuitos se conoce como bibliodiversidad, que contempla la diversidad de los textos que existen, de las modalidades de comunicarlo, de los idiomas, de temas y, en particular, del vínculo que pueden tener muchas de estas revistas en el ámbito local.

PKP busca generar un circuito muy virtuoso en donde las revistas dispongan de herramientas para evidenciar sus trabajos y que logren ser valoradas y, a su vez, consigan más recursos para poder seguir aumentando su alcance. La recomendación para poder obtener el máximo alcance de las alianzas y desarrollos de OJS es tener las últimas actualizaciones y utilizar las nuevas versiones del sistema, que son OJS 3 y OJS 3.4.

Entre las principales razones para llevar a cabo tales actualizaciones están la seguridad, y una serie de funciones de las que mejorarían la calidad de apertura de las revistas y permitiría que las revistas estén mejor integradas con la literatura global y las distintas integraciones.

La versión OJS 3 incluye, en términos de integración, mejoras con ORCID; mejoras con los DOI; disponibilidad de estadísticas; habilidad de exportar usuarios por distintos roles (gestores, editores, coordinadores, correctores, autor, traductor, etc). Para el futuro se contemplan factores como la transparencia, la implementación de revisión por pares abierta en distintas modalidades, la búsqueda de contribuciones y la posibilidad de votar en las opciones del foro.

[Ver conferencia en YouTube](#)

# Evaluación responsable de la investigación en la transición hacia la ciencia abierta en América Latina y el Caribe. CLACSO-FOLEC

CONFERENCIA MAGISTRAL

3 DE OCTUBRE DE 2022

## LAURA ROVELLI

Universidad Nacional de La Plata (UNLP) (Argentina)

### RESUMEN EXTENDIDO

La conferencia propone abordar cuatro objetivos principales: a) el panorama regional a partir de la pregunta ¿por qué importa reformar la evaluación de la investigación para promover la ciencia abierta?; b) algunas iniciativas internacionales y regionales de reforma de evaluación y ciencia abierta; c) recomendaciones de reforma en la evaluación de ciencia abierta, relacionadas con el financiamiento, instituciones, trayectorias, repositorios, y finalmente, d) monitoreo y aprendizajes institucionales.

En cuanto al panorama regional de América Latina y el Caribe, hasta el momento es la región con el mayor porcentaje de sus revistas en acceso abierto, el 25 % de las 29.000 revistas diamante se publican en América Latina, y un 60 % de las revistas diamante se clasifican en las ciencias sociales y humanidades. Las universidades son el principal foco en la puesta en marcha de portales y repositorios de revistas, y recientemente de algunos repositorios de preprints y datos de investigación. Además, América Latina es también una región con amplia coproducción de conocimiento y de contribuciones de actoras y actores sociales fuera de la academia, por ejemplo, en la investigación social, en la investigación de la salud y el medio ambiente.

Bajo este panorama, ¿por qué importa reformar la evaluación de la investigación para promover la ciencia abierta?

A partir de la declaración de Budapest, que en el 2022 cumplió 20 años, empiezan a proliferar un conjunto de iniciativas que diagnostican la necesidad de reformar los sistemas de evaluación y, a su vez, que promueven e impulsan de manera mucho más amplia la ciencia abierta y llaman a modificar los sistemas de evaluación para ponderar estos principios.

Tales manifiestos hacen un llamado al distanciamiento de la evaluación cuantitativa, basada meramente en el factor de impacto o en el índice H de las revistas publicadas en los circuitos internacionales y comerciales de publicación de la investigación, y proponen un modelo de investigación mucho más abierto, participativo, colaborativo, que utilice metodologías cualitativas, situadas, y que estén alineadas con las agendas locales, nacionales de investigación, entre otros principios.

Así mismo, el 6 de junio 2022, se aprobó la Declaración de principios CLACSO-FOLEC que valora la ciencia con relevancia social; la ciencia abierta y la evaluación cualitativa; la ponderación del trabajo en equipo; la participación de la comunidad científica; los indicadores regionales, nacionales e internacionales; la mayor participación de actoras y actores extra-académicos; el impacto predefinido; el multilingüismo; la revisión por pares; la evaluación evolutiva, participativa y transparente; la equidad de las mujeres y diversidades; la inclusividad en etapas tempranas; las trayectorias integrales y diversas y la evaluación situada.

Con respecto a las recomendaciones de reforma en la evaluación de ciencia abierta, relacionadas con el financiamiento, a partir del proyecto IDRC-CLACSO, se han evidenciado aspectos como una creciente evaluación centrada en el acceso abierto y datos abiertos, el fortalecimiento de sistemas de información curricular, portafolios reorientados a la evaluación, la implementación de planes de datos abiertos con protección de datos personales, acceso a registros reutilizables, y de instrumentos de promoción y capacitación en ciencia abierta

y ciudadanía, así como modalidades de AA en resultados de investigación financiada con fondos públicos. El proyecto se puede consultar en [este enlace](#).

Para el tema de la evaluación en instituciones, se centra en tres aspectos: el cambio en cultura de investigación, la formación de estudiantes de grado y posgrado, docentes, investigadores/as; la implementación de políticas institucionales en acceso abierto y ciencia abierta; e instar en la educación superior el trabajo colaborativo y los recursos educativos abiertos, más visibles y fáciles de compartir y reutilizar.

De forma similar, en función de las carreras y trayectorias, se propone el uso de bases internacionales/regionales/nacionales en acceso abierto (sin APC) sin cobrar por leer y publicar, y que estas contribuciones sean en diversas lenguas y formatos; incentivar el uso de principios FAIR y premiar la utilización de software de código abierto; reconocer y valorar la participación activa de sociedad y usuarias y usuarios en el proceso de investigación y promover la revisión por pares y el reconocimiento de diversas contribuciones e interdisciplinariedad.

En relación al tema de los repositorios, se promueve sobre todo la interoperabilidad entre los sistemas y repositorios institucionales, con enlaces permanentes (del tipo DOI, ORCID, handle, ARK), la interacción con infraestructuras regionales o globales con software libres como Open AIRE y LHarvester y que la producción conlleve a procesos de evaluación y de revisión de contenidos entre pares.

El último objetivo conlleva al planteamiento del diagnóstico y de los lineamientos para las políticas de ciencia abierta, que permitan replantear los aspectos metodológicos, en ámbitos internacionales y nacionales para el tema de monitoreo.

A manera de conclusión, se han generado una serie de aprendizajes institucionales en torno a la diversidad de enfoques de evaluación de la CA, en dialogo con trayectorias en CTI e instrumentos y procesos existentes, esto quiere decir, en base a diseños participativos y ascendentes. Así como la

coordinación de políticas y metodologías hacia principios compartidos de CA, el uso de infraestructuras federadas, compartidas, interoperables, sostenibles, que apoyen bibliodiversidad y multilingüismo, y la relevancia del apoyo financiero a la apertura de publicaciones y datos de investigación sin dejar de lado las necesidades en cuanto a infraestructura y formación de recursos humanos calificados.

[Ver conferencia en YouTube](#)

# Gestión de la calidad y la preservación en repositorios institucionales

CONFERENCIA MAGISTRAL

4 DE OCTUBRE DE 2022

## MARISA RAQUEL DE GIUSTI

Universidad Nacional de La Plata (UNLP) (Argentina)

Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

### RESUMEN EXTENDIDO

La conferencia hace un recorrido por los conceptos de repositorios institucionales y de la preservación digital, considerando que los repositorios institucionales son depositarios de la producción de una institución, cualquiera sea su tipología de acuerdo a lo que la propia institución determine. Y que la preservación digital es el conjunto de estrategias, procesos y técnicas que dan respuesta a los problemas que plantea la conservación de los materiales digitales y de los medios (hardware y software) que se emplean para su almacenamiento y consulta, y que en general está destinada a mantener los objetos digitales y sus características de acceso a largo plazo.

Cuando se piensa en preservación digital, se debe hacer referencia como estándar al Modelo OAIS y la norma ISO 14721: ajustar las funciones del repositorio. Lo que plantea es un modelo abstracto que tiene seis entidades: ingesta, gestión de datos, almacenamiento de archivos, preservación, administración y acceso. Esas seis entidades que van desde ingresar el paquete de información, es decir, el contenido que se desea ingresar al repositorio a través de un autoarchivo o bien por operaciones de la administración, u operaciones automáticas de naturaleza informática que toma la forma de un zip.

Durante el proceso se genera un paquete de preservación que se llama AIP, se realizan funciones como separar todos los elementos que son la información descriptiva y enviarla a esta otra entidad que se llama gestión de datos a fin de coordinar el acceso para cuando alguien solicite una información.

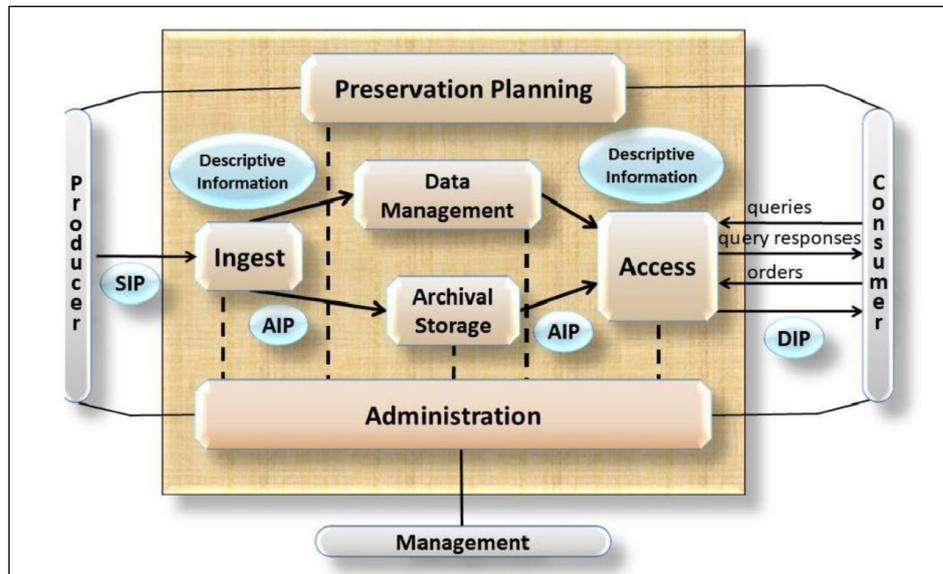


FIGURA 1: OAIS Functional Entities. Management Council of the Consultative Committee for Space Data Systems (CCSDS). (2019). OAIS final v3 draft with changes wrt OAISv2 20190924-rl.docx. P. 41

### El centro de todo: el IP del modelo OAIS

La información descriptiva de la preservación va a tener que contener ciertos elementos, metadatos e información que tiene que ver con la referencia a ese objeto digital y está directamente relacionado con la manera unívoca en que se identifica, por ejemplo, un identificador persistente tipo handle, DOI, etc.

La procedencia (o *provenance*) tiene que ver no solamente con el origen del archivo, sino con los sucesivos tratamientos que se le dieron, es decir, la información con la que está relacionado ese objeto de información. A su vez, *fixity* se refiere a la información de fijeza que asegure que ese archivo no es

alterado; por su parte los derechos se refieren a la licencia para dar difusión a ese contenido, así como la licencia de uso que se le otorgue.

En la imagen siguiente se podrá notar que la información de contenido, si se mira a la izquierda de la figura, se trata de formatos, y a la derecha, se visualizan los metadatos. A su vez, esto se relaciona con tres formas que tomará este paquete de información, mientras se vaya procesando desde la ingesta a las actividades que se realizarán dentro del repositorio, hasta el paquete que se conoce como DIT y que es el paquete que se le entregará a un sistema informático o a una persona que esté solicitando el archivo.

### El IP del modelo OAIS

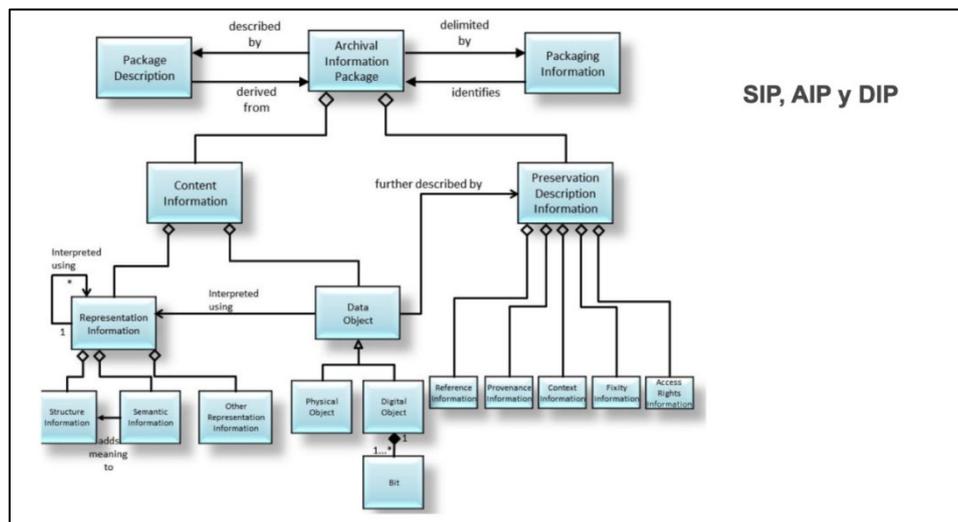


FIGURA 2: Archival Information Package (Detailed View) and its associated Package Description and Packaging Information Management Council of the Consultative Committee for Space Data Systems (CCSDS). (2019). OAIS final v3 draft with changes wrt OAISv2 20190924-rl.docx. P.4-43

El siguiente punto tiene que ver con el trabajo de vigilancia de los contenidos de un repositorio, el cual deberá ejecutarse a lo largo de todo el ciclo, desde que se crea el contenido, se organiza y se clasifica, incluyendo el momento de revisión, en distintas operaciones de verificación e incluso si se elimina el repositorio. Todo ello implica conocer el origen del archivo y si se ejecutaron

otros eventos como migración del archivo original, verificación de su autenticidad por algún tipo de algoritmo, derechos, etc.

### **Preservación de alto nivel**

Lo primero que se debe considerar al hablar de preservación a alto nivel es generar una política de preservación digital; los siguientes puntos se refieren al abordaje sugerido para su realización.

La política de preservación digital de un repositorio debe tener como principal referencia la política institucional de acceso abierto, la cual puede estar en consonancia con la legislación a nivel nacional sobre publicaciones y datos.

Debe establecer la preservación a largo plazo de los contenidos del repositorio.

En lo posible debe hacer referencia a una guía de evaluación de los procesos de preservación en el repositorio (por ejemplo, NDSA), incluso aludir a otros marcos de referencia, por ejemplo, recomendaciones de COAR.

Debe contener un apartado de “Alcances y propósitos”.

Para la creación de un plan de preservación se consideran los siguientes puntos: el plan debe exponer los motivos, principios y sobre qué contenidos va a centrarse para garantizar la conservación, el acceso y la comprensión a largo plazo de esos fondos. Además, se deben identificar necesidades y prioridades y aportar un cronograma pormenorizado que muestre la distribución de las tareas en el período de vigencia del plan.

En cuanto a temas de calidad, para ello es necesario pensar en cómo los repositorios sirven a las personas usuarias, investigadoras, docentes, ¿qué se percibe? ¿qué servicios se brindan? Y en los parámetros e indicadores tanto cualitativos como cuantitativos para objetivar esa percepción de calidad. Considerando las necesidades o expectativas de grupos distintos: cómo son

las diversas comunidades de usuarios y otros sistemas o tecnologías con las que hay que asegurar una interoperabilidad.

### **Forma de evaluación/autoevaluación/revisión**

Sin importar el tamaño o propósito del repositorio, se debe alentar la utilización con una lista de verificación como una herramienta para la evaluación objetiva, interna o externa. La auditoría es la base para comparar las capacidades con un conjunto de criterios centrales para un repositorio digital confiable. La certificación es un paso adicional que algunos repositorios tomarán y/o deberán tomar para el reconocimiento formal y objetivo. El resultado de cualquier auditoría debe verse en el contexto en el que se realizó. Algunos ejemplos de tipos de certificación son DINI (2006) de origen alemán y la ISO 16363.

### **Estructura organizativa**

En términos institucionales, se hace referencia a seis puntos claves a considerar para la creación y el correcto funcionamiento de los repositorios. El primero de estos se refiere a la viabilidad de la organización y su gobierno que considera la misión, el plan estratégico de conservación y la política de colección de los repositorios.

Seguidamente, se plantea la importancia de la estructura organizativa y provisión de personal, cuya base es la identificación de las responsabilidades, la importancia de personal competente, un plan que refleje las funciones de cada parte, y un programa de desarrollo profesional para el personal.

En cuanto al marco de procedimiento de responsabilidad y política de conservación debe tomar en cuenta: la comunidad científica, las políticas de conservación, la historia documentada de los cambios, la transparencia y rendición de responsabilidad, así como las mediciones de integridad de la

información, y finalmente, estar comprometido con un programa regular de autoevaluación y/o certificación externa.

Para el cuarto punto considera la sostenibilidad financiera que involucre la planificación económica a corto y largo plazo, las prácticas financieras y procedimientos transparentes y el análisis e informe de riesgos.

Como quinto y último aspecto de este apartado, se presentan los contratos, licencia y pasivos, que integran los contratos de depósito, la adquisición, mantenimiento y acceso y retirada, las políticas de conservación de contenido, y las políticas de propiedad/derechos.

### **Autoevaluación NSDA**

Para los procesos de autoevaluación una buena práctica es considerar el almacenamiento (copias y localización), la integridad de los datos, la seguridad de la información y los responsables, los metadatos y los formatos.

A su vez existen cuatro niveles para el almacenamiento (copias y localización) y estos tienen una complejidad creciente, estos son: proteger, conocer, controlar y reparar los datos.

Seguidamente, es importante realizar una revisión tanto de prácticas y procedimientos como de contenidos y formatos. Y en cuanto al resguardo de datos, escribir un plan de riesgos que considere la ciberseguridad, fallas de hardware, errores humanos, desastres naturales, etc. El requerimiento de copias de seguridad se define según su cantidad, rotación y ubicación.

Lo anterior se complementa con el abordaje de acciones de preservación en cuanto a integridad (análisis de checksum), el control de cambios (permisos y grupos), los metadatos (uso de identificadores persistentes, formatos estándar, vocabularios controlados y metadatos técnicos), el control de contenido (actualizaciones, migraciones, formatos, virus).

Como conclusiones, la importancia de NDSA es que permite la autoevaluación y porque es relativamente simple. Hay acciones que exceden a

los objetos digitales y la infraestructura, por ejemplo, acciones que hacen a lo organizacional, que no cubre NDSA, básicamente ahí debe escalarse a una norma como la ISO 16363 que además permite certificar el repositorio, pero es bastante compleja.

Como puntos recomendados, se debe definir documentación prioritaria, controlar la ejecución de backups, mejorar la trazabilidad y los metadatos, y valorar la importancia de la política de formatos y de preservación.

[Ver conferencia en YouTube](#)

# Mesa de debate: Experiencias en el desarrollo de Repositorios Nacionales y Regionales de Patrimonio Cultural de América Latina

MESA DE DEBATE

4 DE OCTUBRE DE 2022

## **SANDRA ANGULO MÉNDEZ**

Biblioteca Nacional de Colombia (Colombia)

## **GABRIELA AYRES FERREIRA TERRADA**

Biblioteca Nacional de Brasil (Brasil)

## **DAVID ESQUIVEL PALOMARES**

Oficina para la Memoria Histórica de México (México)

## **CHRISTIAN ZAMORA RODRÍGUEZ**

Universidad de Costa Rica (Costa Rica)

### **OBJETIVO**

Conocer e intercambiar experiencias sobre el desarrollo y sostenibilidad de repositorios nacionales o regionales de patrimonio cultural, de cuatro instituciones representativas de Colombia, Brasil, México y Costa Rica. La mesa de debate plantea una primera ronda que aborda el siguiente interrogante: *¿Cuál es el alcance de su proyecto y en qué consiste?*

### **Biblioteca Nacional Digital (Colombia)**

La Biblioteca Nacional Digital de Colombia con este proyecto evidencia que existe una serie de funciones en una biblioteca digital, que va más allá de una colección digitalizada y de una herramienta de gestión de la información. Bajo estos principios, la BNC se ha centrado en funciones como:

- Recuperar, preservar y dar acceso a la memoria colectiva del país, representada por el patrimonio bibliográfico y documental
- Promocionar y fomentar las bibliotecas públicas
- Planear y diseñar políticas de lectura

Pero, además, ha permitido el acceso abierto a documentos únicos, frágiles y representativos del PByD, y de otras bibliotecas del país, generado contenidos sobre Colombia o escritos por colombianos para la investigación, educación y difusión, producido colecciones en línea y contenidos digitales culturales y elaborado un programa de preservación digital.

### **Biblioteca Nacional de Brasil**

Como menciona Gabriela Ayres Ferreira a partir del Estatuto FBN - Decreto N° 8297, de 15 de agosto de 2014 de Brasil, se establecen como principales funciones de la biblioteca:

- Captar, conservar y difundir los registros de la memoria bibliográfica y documental nacional
- Adoptar las medidas necesarias para la conservación y protección del patrimonio bibliográfico y digital bajo su custodia
- Promover la cooperación y difusión nacional e internacional relacionada con la misión de la Fundación Biblioteca Nacional

Y en función de la digitalización, en el 2000 el crecimiento tecnológico y de la web impulsaron el Laboratorio de Digitalización de la Información, y en el 2006 la creación de la UN digital, y un año más tarde, el primer proyecto de memoria digital, desde la memoria brasileña virtual en la plataforma de Space.

## **MEMORICA (México)**

Tal como comenta David Esquivel Palomares, el proyecto surge en la iniciativa del gobierno del presidente Andrés Manuel López Obrador que comenzó en 2018, el repositorio se presentó en 2020 muy poco antes de la pandemia, coincidiendo con un momento donde este tipo de plataformas o iniciativas en territorio nacional evidenciaron su necesidad. El principal objetivo de MEMORICA es poner a disposición del público archivos digitales de variada naturaleza relacionados con la historia y las expresiones culturales de México, fomentando con ello el derecho a la memoria.

Así mismo, se trabaja con diferentes tipologías documentales y lo que se busca es incluir, de una manera cada vez más amplia el conocimiento de diferentes acervos que resguardan la memoria de México, esto incluye: libros, fotografías, revistas, mapas, obra artística, objetos prehispánicos, entre otros.

La iniciativa no cuenta con un archivo propio bajo custodia, sino que responde a una necesidad inter-institucional de diversas entidades, tanto públicas como privadas y trabaja con fototecas estatales o municipales de todo el territorio mexicano, no solamente se suscribe a una parte federal, a pesar de que el proyecto está inscrito en el aparato federal gubernamental, sino que también se trabaja con todos los órdenes de gobierno estatal municipal. Además de colecciones privadas de grandes empresas como la Fundación Carlos Slim, a través del Centro de Estudios Históricos de México (CASO), y de personas particulares que han ofrecido que sea MEMORICA la ventana de difusión de materiales que de otra manera no se podrían conocer.

## **Repositorio Centroamericano de Patrimonio Cultural (Costa Rica)**

En este caso, Christina Zamora comenta que se trata de un espacio para la difusión de colecciones especializadas en patrimonio cultural, tanto costarricense como centroamericano, que se gesta desde 2013, y que fue impulsado desde la Universidad de Costa Rica por medio del Instituto de Investigaciones en Arte (IIARTE).

El proyecto tiene por objetivo visibilizar las investigaciones inter y transdisciplinarias vinculadas a las artes, para fomentar el acceso libre a bases de datos de patrimonio cultural y para ello pone a disposición más de 7.000 obras curadas por colecciones de diversas características como archivos sonoros, artes gráficas, arte moderno, fotografía, moda, entre otros.

El repositorio cosecha colecciones de patrimonio cultural de centros de investigación y bibliotecas a nivel centroamericano quienes creen en la visión de este proyecto: acceso libre al patrimonio de la región centroamericana (Open GLAMA). En 2021, democratizar el conocimiento sobre la cultura en la región permitió su ingreso al Registro Nacional de Memoria del Mundo de la UNESCO. Actualmente, el repositorio recibe más de 5 millones de visitas, la mayoría de países se ubican fuera de Latinoamérica, esto sucede en especial con países como Estados Unidos, Japón, Italia, Emiratos Árabes, entre otros. Esto es un indicio de que el patrimonio cultural, visual y sonoro de la región centroamericana está siendo investigado o es de interés para personas de contextos muy lejanos al centroamericano.

Segunda ronda: Desde el enfoque del modelo GLAMUR (Gallery, Library, Archive, Museum, University & Repository), *¿qué estrategias han implementado para la digitalización, preservación y doblamiento de colecciones documentales y de arte? y ¿cuáles son las perspectivas de desarrollo de su repositorio nacional/regional?*

### **Biblioteca Nacional de Colombia**

El proyecto de la biblioteca nacional inició con alrededor de 3.400.000 y a partir de un trabajo de selección esta colección se redujo a 250.000 objetos que entraron al proceso de digitalización para el acceso abierto. Siendo esta la primera estrategia de acceso abierto, cuenta con una serie de libros, material hemerográfico, fotografías, material audiovisual, proyectos de cine experimental, entre otros.

Además de las 27 exposiciones de la biblioteca digital y con base en todas las colecciones del repositorio analógico, existen también proyectos como la Gaboteca que hace homenaje a Gabriel García Márquez, proyectos como la Biblioteca Básica de Cultura Colombiana, que recoge las principales obras tanto de clásicos como de libros colombianos, o la Biblioteca Digital Familiar, que a través de un código QR permite el acceso a alrededor de 146 libros.

### **Biblioteca Nacional de Brasil**

Desde 1994, la Biblioteca Nacional de Brasil ha llevado a cabo actividades sistemáticas de transferencia de soporte como la microfilmación y la digitalización, con el objetivo de difundir y preservar las colecciones bajo su custodia, y con el fin de orientar el proceso de selección de las obras que lo integran, se establecieron dos criterios que toman en cuenta aspectos como calidad y dominio público.

El trabajo se realiza en conjunto con las áreas de colecciones, es decir, con los curadores para hacer la digitalización, de esta forma, la Biblioteca Nacional posee actualmente 14 millones de documentos denominados por el Programa Memoria del Mundo. La digitalización sistemática con la preservación a largo plazo de diversas colecciones hace que se pueda disponer de un gran acervo en la web. Actualmente se está trabajando en un plan de preservación digital para la sustentabilidad del programa.

### **MEMORICA (México)**

Con respecto a la implementación del acrónimo GLAMUR, David resalta la importancia de entender la diversidad de entidades que conforman el universo GLAMUR, porque cada una de las entidades responde a una práctica documental distinta, la práctica documental del archivo, no es la misma que la práctica documental de la biblioteca, o de la fototeca o del museo. Sus naturalezas son muy diferentes, no solamente con respecto a la administración

de los recursos, sino a la custodia, la conservación, y por ende su preservación, pero sobre todo a nivel de la circulación.

La circulación de la obra artística, por ejemplo, es completamente distinta en sus términos más amplios a las estrategias que pudiera llevar a cabo una biblioteca, entre otras razones, por la propia naturaleza del objeto, pero también por una cuestión legal y es ahí donde la iniciativa Open GLAM busca generar las condiciones de acceso a materiales que no necesariamente cuentan con los derechos.

Otro aspecto importante tiene que ver con cómo las instituciones van generando la calidad de la información, es decir, contar con ciertas licencias, como Creative Commons o la iniciativa del Repositorio de Europeana, además de cómo se genera cada vez mejor metadata de información disponible al usuario, si bien, se sabe que la mayoría de la metadata existente al interior de las entidades está relacionada con el control administrativo que se podría llevar a cabo al interior de cada entidad, en ocasiones no se integra más allá del espacio específico, otros tipos de conocimiento que surgen fuera de la propia institución.

Así mismo, MEMORICA ha tenido la fortuna no solamente de integrar colecciones correspondientes a la memoria de México a nivel nacional, sino también a nivel internacional, como colecciones de Gálica, de la Biblioteca Nacional de Austria, del Instituto Cultural Prusiano, que es la red más importante de entidades de custodia de memoria que existe en Alemania. Estos ejemplos apuntan a la importancia de generar proyectos que se enfoquen en la mejora de la calidad de la información, más allá del mero control administrativo, permitiría identificar cuestiones como el periodo, agente creador, fechas específicas, o información para el usuario. Lo importante es generar las condiciones de acceso colectivo que se separen de discursos hegemónicos que han promovido algunas instituciones y que dejan de lado el pensamiento colectivo de lo que representa la memoria para un país. Y la importancia generar las condiciones de acceso que promuevan procesos de construcción de memoria propios, comunitarios que no necesariamente estén

circunscritos a una administración, o una entidad, sino que sean más horizontales, para que ese diálogo pueda ser cada vez más amplio.

De esta forma, la iniciativa busca generar políticas de digitalización más robustas, considerando que un proyecto de digitalización debe considerar temas de planeación, sobre todo cuando se habla de preservación digital y metadata. El otro punto tiene que ver con generar colecciones digitales sustentables, donde se cuestione si el material que se desea digitalizar no ha sido digitalizado previamente, o de la eficiencia en el gasto de recursos públicos. Y, finalmente, promover que no haya tecnología propietaria al interior de las entidades públicas, sino que, se apueste por el desarrollo tecnológico propio con el talento de las personas que puedan conceptualizar esos sistemas, esto para evitar la posibilidad de que las entidades dependan en gran medida de grandes monopolios o que apuesten por bibliotecas o iniciativas libres, siempre bajo un cuestionamiento crítico sobre las necesidades de la iniciativa en cuestión, para seleccionar una infraestructura más acorde con estos requerimientos y apostar por una visión de gobernanza más sostenible.

### **Repositorio Centroamericano de Patrimonio Cultural (Costa Rica)**

El repositorio ha trabajado con diferentes variedades de actores para poder digitalizar y preservar las diferentes obras de patrimonio cultural, además, apuesta por la adopción de un esquema de datos que le permita realizar una descripción, que no se limite únicamente a especialistas, sino también a las personas participantes dentro de procesos propios de preservación, restauración y conservación, y que estos puedan comprender cómo aprovechar ese tipo de información para la búsqueda.

Algunas de las colecciones, por ejemplo, desde la Biblioteca Arturo Agüero Chávez, en Costa Rica, se realiza un proceso en donde la digitalización de la obra incluye además ciertas características que se pueden aprovechar para obtener metadata de suma importancia para la investigación y la difusión, y enfatizando que se aborde pensando en el usuario final, haciéndolo de la forma

más heterogénea posible y con términos que le permitan una búsqueda sencilla. Y definitivamente para la creación del repositorio, ha sido muy importante el factor tecnológico que ha permitido desde el Instituto de Investigaciones de Arte mantener la preservación digital de la información.

En relación a la preservación, las instituciones por el contexto económico de cada una de ellas, no cuenta con un sistema de preservación y almacenamiento, por tanto, se han desarrollado en lo interno procedimientos para el manejo de formatos que permitan el mantenimiento a largo plazo, y, además, afrontar las condiciones a las que Centroamérica está expuesta como región, considerando elementos de tipo ambiental y tecnológicos.

*Tercera ronda: ¿Cuáles han sido sus principales retos, fortalezas y aprendizajes en el desarrollo del proyecto, así como los beneficios para sus comunidades y la sociedad?*

### **Biblioteca Nacional de Colombia**

En el caso de la Biblioteca Nacional de Colombia, el proyecto se planeó en una etapa preliminar con tecnología analógica (microfilmación) y eso era lo más cercano al proceso digital, entonces, ese nicho de formación fue importante e interesante para poder planificar un proyecto considerando las dificultades económicas, y la obsolescencia de la tecnología.

Concretamente en el tema de fortalezas y retos, se optimizó todo el proceso de catalogación porque se hizo uso de una tecnología propietaria, pero al mismo tiempo se desarrolló un «híbrido» que integra la página web con el repositorio. Todo entra en el repositorio digital con el objetivo de buscar la preservación a largo plazo, esto con el fin de implementar metadatos que se consideran de gran importancia. Así mismo, se ha desarrollado el flujo de trabajo a partir de las directrices y estructura de la biblioteca nacional digital, la misma política de preservación, la sostenibilidad y la formación del equipo.

Lo anterior constituye un proceso muy importante porque en el caso del patrimonio bibliográfico y documental frágil de la biblioteca, se requiere tener antecedentes y conocimientos sobre la conservación de los objetos para el proceso digital. Se rescata además el proceso de aprendizaje sobre la normalización de los formatos digitales para precisamente combatir la obsolescencia tecnológica. Esto resulta fundamental, considerando que el repositorio fue creado desde la perspectiva de la preservación en términos de acceso abierto y democrático a bienes culturales.

En cuanto a los retos, destaca principalmente el tema de una infraestructura tecnológica activa. Esto en tanto se requiere de un resguardo del datacenter y la metadata, además del software, el hardware, las conexiones con la tecnología propietaria, los desarrollos internos etc., frente a lo cual se ha desarrollado un Plan Nacional de Patrimonio Bibliográfico y Documental, que se llama «Vamos a hacer memoria», que pretende ampliar los contenidos tanto locales como regionales y nacionales para poder dar cuenta de esas distintas voces que conforman el patrimonio de la nación. Esto implica para Colombia un nuevo momento histórico a través de su plan de gobierno, y de su enfoque de Cultura y Paz, en tanto la colección digital esté alineada con la justicia social, ambiental, económica, y en el caso de la biblioteca hacerlo a través del repositorio.

### **Biblioteca Nacional de Brasil**

En el caso de Brasil se presentan desafíos en torno a una mayor inversión para el mantenimiento y ampliación del servidor de la BN digital con el fin de garantizar la calidad y la disponibilidad constante al servicio de la población. Aunado a esto, está la ampliación de la infraestructura tanto del laboratorio de digitalización como del sector de procesamiento técnico, además de la ampliación del espacio de almacenamiento para garantizar la seguridad y la sustentabilidad del programa.

En cuanto a los beneficios, puede destacarse que el trabajo continuo durante la pandemia impulsó el crecimiento de los accesos y el reconocimiento de la importancia del programa de la BN Digital para las personas investigadoras y para la población brasileña, así como la de otros países. En cuanto a las expectativas se pretende implementar la política de preservación digital, el desarrollo de un plan de preservación y ampliar la capacitación de técnicos, así como aunar esfuerzos de inversiones para la mejora de los servicios de la biblioteca digital.

### **MEMORICA (México)**

El principal reto consiste en seguir integrando la diversidad en todas las tipologías documentales que se manejan, incluyendo la variedad de formas prácticas documentales, y formas de catalogación que existen al interior de México, entendiendo que la cultura documental y la cultura de catalogación existente es muy variada. La integración de todo ese conjunto de información bajo un mismo modelo de datos, por muy flexible que sea, así como el procesamiento de la base de datos, constituye uno de los retos más importantes, porque si bien existe una guía muy específica de normatividad siempre hay un nuevo reto que atender. Existe también la necesidad de continuar sensibilizando a las entidades, con respecto al tema de tener acceso abierto y seguir mejorando el modelo de datos, y la estructura tecnológica que soporta todos estos materiales.

En cuanto a las perspectivas, una es publicar parte de los catálogos, en los que se usan vocabularios controlados, no solamente para la aplicación del guardado de la información sino para que puedan ser recuperables semánticamente. La otra tiene que ver con el trabajo de validación y la formación de catálogos para que pueda ser de utilidad a otros proyectos.

Finalmente, se desea escalar a un repositorio semántico que permita una interacción mayor con otros proyectos y seguir generando una política para la

recuperación de la memoria artística biocultural de México, existente también fuera de nuestro país.

### **Repositorio Centroamericano de Patrimonio Cultural (Costa Rica)**

Para el repositorio, uno de los mayores retos es seguir haciendo comunidad en Centroamérica, se desea que por medio del repositorio se puedan integrar otras regiones. Se han realizado durante el último año algunos intercambios de diferentes tipos de obras de arte.

Además de esto, se considera el manejo de los sets de los datos y la preservación con un repositorio de datos enlazado, que permita compartir estos sets y que sean completamente transparentes por medio del repositorio.

**[Ver mesa de debate en YouTube](#)**

# Gestión de Recursos Educativos Abiertos y Prácticas Educativas Abiertas

## Recomendaciones UNESCO 2019

CONFERENCIA MAGISTRAL

5 DE OCTUBRE DE 2022

**DIANA HERNÁNDEZ**

FAB LAB Ká Träre, Universidad Estatal a Distancia (Costa Rica)

### RESUMEN EXTENDIDO

La conferencia inicia con la definición de los Recursos Educativos Abiertos (REA), los cuales son materiales de aprendizaje, enseñanza e investigación, en cualquier formato y soporte, de dominio público o protegidos por derechos de autor y que han sido publicados con una licencia abierta que permite el acceso a ellos, así como su reutilización, reconversión, adaptación y redistribución sin costo alguno por parte de terceros.

Por su parte, las Prácticas Educativas Abiertas (PEA), según el International Council for Open and Distance Education (ICDE), son «las prácticas que apoyan la producción, utilización y reutilización de Recursos Educativos Abiertos de alta calidad, por medio de políticas institucionales, promoción de modelos pedagógicos innovadores y el respeto y el empoderamiento de los estudiantes como coproductores de sus aprendizajes a lo largo de la vida».<sup>1</sup>

---

<sup>1</sup> International Council for open and Distance Education (ICDE) (2011). Definition of Open Educational Practices. Retrieved October 17th, 2016, from [www.icde.org/en/resources/open\\_educational\\_quality\\_initiative/definition\\_of\\_open\\_educational\\_practices/](http://www.icde.org/en/resources/open_educational_quality_initiative/definition_of_open_educational_practices/)

Se abordan a continuación cinco ámbitos de acción de REA según la experiencia del OpenCourseWare de la Universidad Estatal a Distancia (UNED) de Costa Rica y Open Education Global.

En el ámbito de acción, sobre el desarrollo de capacitaciones de las partes interesadas en materia de creación, acceso, reutilización, adaptación y redistribución de REA. Se hace alusión al Nodo REA, como una forma de gestión para trabajar el tema y mantenerlo presente en la institución, a partir de un taller de Recursos Educativos Abiertos y Prácticas Educativas Abiertas. Para esto se abordaron temas como las licencias abiertas y la adaptación de materiales para compartir con otras personas. Actualmente, no solo se comparten cursos, sino también otros recursos como sistematizaciones de investigaciones que se han realizado, diseños o modelados 3D, y otros elementos, lo cual amplía la visión sobre qué son los REA. En el nodo hay representantes de todas las escuelas y también del área de investigación. Para el caso de Open Education Global, un ejemplo es el desarrollo de un podcast donde personas de diferentes instituciones y diferentes áreas hablan sobre iniciativas relacionadas con estos temas.

En relación con el segundo ámbito, que tiene que ver con la elaboración de políticas de apoyo, en la UNED se han realizado acuerdos como el del Consejo de Rectoría apoyando la membresía en el OE Global, OCW, UNED y las licencias abiertas, además de la inscripción del proyecto que sustenta el Nodo REA. También la declaratoria de interés institucional de estas Recomendaciones de la UNESCO y la propuesta de una ponencia en el V Congreso Universitario que tiene como objetivo ampliar la flexibilidad curricular y el acceso a la educación abierta desde la producción institucional de Recursos y Prácticas Educativas Abiertas.

Es así como en la UNED se ha propiciado la idea de trabajar en abierto, más allá del trabajo que se pueda hacer con los estudiantes, se busca aplicar a proyectos de investigación. Por ejemplo, existe un reconocimiento para quienes trabajan con dispositivos tecnológicos abiertos. Mientras que en OE Global se

tiene un plan estratégico 2021-2030 que marca opciones en las que se pueda trabajar y desarrollar proyectos a nivel de políticas en cada institución.

En el tercer ámbito que se refiere a la promoción del acceso efectivo, inclusivo y equitativo a REA de calidad. Estas incluirían a los educandos en contextos de educación formal y no formal, independientemente, entre otras cosas, de su edad, género, capacidad física y situación socioeconómica, así como a quienes se encuentran en situación de vulnerabilidad, los pueblos indígenas, las personas en zonas rurales remotas, quienes viven en zonas afectadas por conflictos y desastres naturales, las minorías étnicas, los migrantes, los refugiados y los desplazados.

Como ejemplo, en la UNED existe una serie de unidades didácticas que integran recursos educativos abiertos, que se pueden visualizar en la siguiente imagen:



Así mismo, el Ministerio de Educación Pública de Costa Rica generó un compendio REA para maestros, estudiantes y padres, el cual se trabajó desde un nodo. Por su parte, en OE Global, en marzo de 2023 se realizará la Semana de la Educación Abierta, y se cuenta con los premios del Open Education Global.

Como cuarto ámbito se plantea el fomento de la creación de modelos de sostenibilidad para los REA. Esto quiere decir, la movilización de recursos no tradicionales basada en la reciprocidad y mediante alianzas y establecimiento de redes, la generación de ingresos, como donaciones, afiliaciones, pago del importe que se desee y financiación colectiva, que pueden proporcionar ingresos y sostenibilidad a los REA. En este sentido se ha trabajado en la cooperación interuniversitaria para la formación e investigación en el movimiento educativo abierto (2014-2016) Tecnológico de Monterrey.

Finalmente, el quinto y último ámbito tiene que ver con la promoción y fortalecimiento de la cooperación internacional. Promoviendo el fortalecimiento de capacidades, repositorios, comunidades de prácticas, investigación conjunta sobre REA y solidaridad entre todos los países, independiente del nivel de desarrollo de sus REA.

A nivel de la UNED y aliados OE Global existe OELATAM, Educación Abierta en Latinoamérica, que integra el estado del arte a nivel de cada país, y formación docente (licencias abiertas).

[Ver conferencia en YouTube](#)

# ¿Qué es la Research Data Alliance?

CONFERENCIA MAGISTRAL

6 DE OCTUBRE DE 2022

## REBECCA KOSKELA

Research Data Alliance (Estados Unidos)

### RESUMEN EXTENDIDO

Para abordar la temática es necesario identificar que la Research Data Alliance (RDA) es una organización global que dio inicio en marzo 2013 por iniciativa de la Comisión Europea, Estados Unidos y la Fundación de Ciencias de Estados Unidos y el gobierno de innovación de Australia, con el propósito de construir los puentes técnicos que permiten el compartir el uso de datos.

Existe, además, una serie de principios que guían su accionar. El primero de ellos es apertura, es decir, que está abierto a cualquier individuo interesado u organizaciones que se suscriben a esos principios. Las reuniones de la comunidad y otros procesos están abiertas y los entregables y aportes de la RDA son diseminados públicamente.

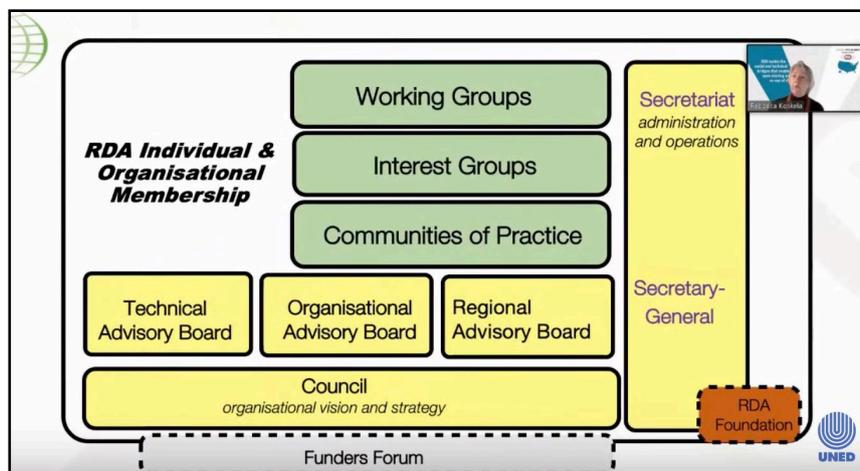
La segunda característica es consenso, en cuanto a la resolución de cualquier tipo de conflicto. Así mismo, se apuesta por el balance, a partir de una representación equilibrada de sus miembros y comunidades de partes interesadas, la armonización en cuanto a normas de datos, políticas y tecnologías, su infraestructura y comunidades. Otro aspecto que tiene peso es la representación comunitaria, con el Community-Drive, en tanto es liderado por las comunidades, constituido por voluntarios miembros y organizaciones, y todos son apoyados por el secretariado de la RDA y, finalmente, la RDA apela por Non-Profit, ya que no promueve y tampoco endosa, o vende productos comerciales, tecnologías o servicios.

La siguiente imagen muestra cómo ha ido creciendo RDA a lo largo del mundo, actualmente hay más de 12.000 miembros de 148 países. La mitad de

miembros proviene de Europa y el segundo lugar de Norteamérica, seguido por Asia, Oceanía, África y Latinoamérica. Hoy en día Latinoamérica constituye el 6 % de la membresía, La mayoría de miembros de RDA son investigadores académicos, seguido de personas de administración pública, la industria, la prensa y sociedades.



El siguiente es un organigrama que muestra cómo se compone la RDA.



Tal y como se puede visualizar en la imagen, los tres grupos principales se componen de grupos de trabajo, grupos de interés y las comunidades de prácticas. En el lado de negocios se compone principalmente por una junta asesora técnica, una junta asesora organizacional y una junta asesora regional.

Posteriormente, está el consejo que es el responsable de la visión y de mantener la revisión y la estrategia de la RDA. Las personas que hacen la gran mayoría de trabajo en el apoyo de grupos son las que conforman el Secretariado General. Finalmente, la Asamblea Organizacional se compone de 70 organizaciones, algunas de estas son afiliados, por ejemplo, ORCID, etc.

En cuanto al tema de las regiones, también existen diferentes tipos con sus características propias, la primera son las regiones interesadas, que son aquellas que están conociendo la RDA, pero quieren estar un poco más involucrados. Seguidas de las regiones participativas, las cuales tienen un lugar en el sitio web, pero no hay un memorando de entendimiento establecido. Luego, están las regiones comprometidas, que son las que han avanzado más en su relación con la RDA y tienen un memorando de entendimiento, pero no hay ningún tipo de apoyo financiero o de apoyo en especie. Y finalmente, las regiones contribuyentes, que también tienen un lugar en el sitio web y ofrecen apoyo en especie o financiero.

<b>LISTADO DE REGIONES QUE CONFORMAN RDA</b>			
<i>Interesadas</i>	<i>Participativas</i>	<i>Comprometidas</i>	<i>Contribuyentes</i>
Suiza	Américas Austria Brasil Bulgaria Croacia Republica Checa Dinamarca Estonia Finlandia Alemania Grecia Hungría Irlanda Italia Lituania Holanda Noruega Rumania Eslovenia España Suecia Reino Unido	Costa Rica Portugal	Australia Canadá Europa Francia Estados Unidos

El siguiente aspecto que aborda la conferencia es ¿qué hace la RDA?

Los miembros se reúnen y se forman los grupos de interés o los grupos de trabajo y abordan desafíos que son compartidos entre las personas del grupo global. Lo que hace a la RDA diferente de muchas organizaciones es que existe la posibilidad de poder trabajar con personas alrededor del mundo que tienen problemas muy similares y que han logrado averiguar cómo resolver ese

problema para después compartir consejos, es decir, se puede trabajar en un problema conjuntamente.

Algunas de las áreas en particular que se están trabajando son: reproductibilidad de la responsabilidad y los resultados, servicios de vocabulario, mejores prácticas o repositorios, interoperabilidad legal, citación de datos, registros de tipos de datos y metadata.

Se puede observar desde el punto de vista del ciclo de vida de gestión de datos. En tanto hay grupos que trabajan en planes (por disciplinas, por planes de gestión de datos automatizados), además, existen grupos que están viendo cómo se descubren los datos, estos diferentes grupos son accesibles y operativos.

Así mismo, se cuenta con un área para ambientes de investigación de datos, como el grupo de trabajo de software de investigación y luego la sección de archivos del ciclo de vida de gestión de datos y que relaciona su trabajo con repositorios y dominios.

En relación a las publicaciones, están quienes trabajan en las diferentes editoriales, lo hacen para tratar de alinear las propuestas en estos términos. Posteriormente, está la planificación, que es un área bastante grande, y tiene que ver con ética y con los principios de atención que involucran, por ejemplo, datos indígenas y de asegurar la correcta citación y acreditación sobre de dónde provienen esos datos y personas responsables de los mismos. Finalmente, existen un grupo de inteligencia artificial y de citación.

El siguiente aspecto que aborda la expositora es el de las recomendaciones y productos, y esto es que existen productos de apoyo, que son básicamente soluciones útiles que vienen de los grupos de trabajo o los grupos de interés. Los principales enfoques en las recomendaciones son para procesos de gestión de datos, mejores prácticas, sobre todo en cuanto a recuperación de datos, herramientas para los datos, descripción de datos para permitir su reutilización, y finalmente, la categoría de identidad, almacenamiento y preservación de los datos.

La filosofía de los grupos de trabajo básicamente es crear una solución y que sea adoptada y utilizada por varias organizaciones. Esto básicamente proviene de los grupos de trabajo en este momento se tienen 64 recomendaciones principales y productos y más de 200 casos de adopción en varios países y organizaciones. Para más información, se puede acceder [este enlace](#).

La última sección que se aborda en la conferencia es la de los pasos para involucrarse y participar en la RDA. Una vez que se es miembro, se puede unir a un grupo existente o comenzar un nuevo grupo, y en estos últimos existen tres categorías: los pájaros de una pluma (expresan su problema para que las personas puedan manifestar su interés en participar en su resolución), además están los ya mencionados grupos de interés, y los grupos de trabajo.

[Ver conferencia en YouTube](#)

# Consortia Papyrus – Repositorio de Datos de Investigación en Colombia, una iniciativa de trabajo colaborativo

CONFERENCIA MAGISTRAL

6 DE OCTUBRE DE 2022

## CÉSAR AUGUSTO RENDÓN VALENCIA

Consortia (Colombia)

### RESUMEN EXTENDIDO

El experto presenta el proyecto Consortia Papyrus a través de una serie de preguntas, a partir de las cuales se plantean las respuestas.

#### ¿Qué es un repositorio de datos?

Un repositorio de datos es una herramienta a través de la cual académicos e investigadores almacenan y realizan una amplia descripción de los datos producto de su ejercicio investigativo, con el fin de divulgarlos, gestionarlos y facilitar el acceso y uso por parte de la comunidad científica y el público en general.

#### ¿Por qué una institución debe gestionar los datos de investigación?

Existen varias razones por las cuales realizar esta gestión, una de ellas es la exigencia de las revistas, hoy en día es obligatorio llegar a acuerdos sobre en dónde estarán publicados los datos que sustentan los estudios, ya no basta con que estén en los propios computadores de las personas docentes o investigadoras, sino que tienen que estar visibles; además para una institución de educación superior es muy relevante la posibilidad de tener visibilidad.

Del mismo modo, las agencias financiadoras están exigiendo que estos datos sean públicos y que los datos se puedan consultar para poder hacer desarrollos futuros. Y así mismo, las propias instituciones y las vicerrectorías de investigación están solicitando que se realice una gestión centralizada de los datos que se preserven en el tiempo, además de facilitar la recuperación y la reutilización de los mismos aporta transparencia a los procesos de investigación que estamos realizando.

### **¿Cuál es la principal motivación tras la creación de Consortia Papyrus – Repositorio de Datos de Investigación?**

En primer lugar, está la necesidad de fortalecer el trabajo colaborativo, además de impulsar la ciencia abierta en el país y, de esta forma, visibilizar la producción nacional en escenarios internacionales.

Por otra parte, está la tarea de disminuir costos en todos los aspectos técnicos, considerando el contexto latinoamericano. En términos de un proyecto tecnológico cuando estamos hablando de desarrollos a la medida o de instalaciones, servidores, espacios en disco, etc., se hace evidente la necesidad de colaborar, y finalmente aprovechar infraestructura tecnológica y recurso humano calificado.

### **¿Qué es Consortia?**

Se trata del operador del consorcio de Colombia que reúne 61 instituciones entre públicas y privadas, y en el marco de este proyecto de consorcio, se han generado un sinnúmero de actividades y de subproyectos. Aunado a esto, el consorcio ha apoyado el desarrollo de ciencia abierta y realizado seminarios, empezando por el seminario de acceso abierto.

El siguiente punto tiene que ver con la arquitectura, que es el modelo de funcionamiento de Consortia. Así es como se creó un diseño que tiene dos tipos de espacio, uno es el espacio de gestión conjunta y el otro un espacio de

recolección. Este nivel de arquitectura permite que, cuando una institución no tiene aún desarrollado su repositorio de datos y no está trabajando aún con Dataverse, de igual forma tiene la posibilidad de participar en el espacio de gestión conjunta en donde se le creó su propio Dataverse para poder desarrollarlo con las características requiera.

De esta forma, se comparte servidor, espacio en disco y una serie de elementos que se unifican en el repositorio. Pero también se respeta el espacio de recolección de aquellas instituciones que tienen desarrollos propios que ya tenían en el pasado instalado un Dataverse en funcionamiento, pero que querían participar en el consorcio. Para esto se creó un espacio de recolección en donde se recolectan los datasets que ya tienen a disposición y así pueden participar del proyecto conjunto.

A continuación, se describen algunas características propias de Consortia:

**Definiciones institucionales:** estas definiciones se toman en función de una serie de variables, por ejemplo, la distribución de los Dataverse. Es decir, es muy importante conservar la estructura que tiene cada institución en términos de la investigación, entonces cada institución puede decidir dentro de su Dataverse propio cómo se va a estructurar.

**Tipos de datos:** actualmente, la plataforma permite datos de tipo textual, multimedia, imágenes, tabulares, estructurados, código fuente, específico de programas y archivos netCDF.

**Licencias de uso:**

**Abierto:** contenido visible por todo usuario de internet. Su uso puede estar sujeto a distintas limitaciones según la licencia Creative Commons con la que se registre.

**Cerrado:** contenido visible sólo para el personal encargado de la gestión del repositorio institucional.

**Restringido:** contenido en principio cerrado, pero que puede ser puesto a disposición del público bajo ciertas condiciones y siempre bajo petición.

**Embargado:** contenido cerrado que pasa automáticamente a ser abierto, una vez transcurrido un período de tiempo determinado.

**DOI en Papyrus (repositorios de datos):** a cada organización miembro del consorcio se le asigna un prefijo diferente, que permitirá identificar a qué institución pertenece cada DOI asignado. Los DOI pueden ser asignados a sets de datos en el repositorio de datos de investigación o a otros resultados de investigación que estén almacenados en el repositorio institucional como tesis, trabajos de grado, preprints, entre otros. Cada organización tendrá hasta 10.000 DOI para asignar.

### **¿Cuáles han sido los alcances de Consortia Papyrus?**

El proyecto, que se inició en enero 2022, desde su comienzo se dio a la tarea de adquirir un servidor, un dominio, y elaborar una interfaz gráfica para la instalación de Dataverse en el servidor, así como la creación de los espacios para cada institución miembro.

Hasta ese momento la Universidad del Rosario participó en el proceso de integración de las otras 12 instituciones, actualmente participan 13 instituciones que están en el espacio de gestión conjunta, cada una con su propio espacio. Por otra parte, se han realizado distintos Webinar en temas como: capacitación técnica, plan de gestión de datos, definición de metadata, variables transversales, y la gestión de datos en la investigación.

**[Ver conferencia en YouTube](#)**

# Políticas institucionales de gestión de datos: buenas prácticas, retos y casos en Brasil, Colombia y Argentina

MESA DE DEBATE

6 DE OCTUBRE DE 2022

## MARISA RAQUEL DE GIUSTI

Universidad Nacional de La Plata (UNLP) (Argentina)

Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

## MALGORZATA LISOWSKA

Universidad del Rosario (Colombia)

## CATERINA GROPOSO PAVÃO

Universidade Federal do Rio Grande do Sul (Brasil)

### RESUMEN EXTENDIDO

**CATERINA G. PAVÃO** (Universidade Federal do Rio Grande do Sul, Brasil):

Desde el 2001, se han realizado esfuerzos a partir del Instituto Brasileño de Información en Ciencia y Tecnología (IBICT) con la creación del programa de Biblioteca Digital Brasileña, englobando el proyecto de Biblioteca Digital de Tesis y Disertaciones.

En el 2021 se realizó la creación de Redes Brasileñas de Repositorios Digitales formadas por las Redes Regionales de Repositorios Digitales (RBRD).

A continuación, se sintetizan algunos de los principales avances en cuanto a legislación en repositorios de datos y ciencia abierta:

- Ley de Acceso a la Información - LAI (Ley N° 12.527/2011): regula el derecho constitucional de los ciudadanos a acceder a la información pública y es aplicable a los tres poderes de la Unión.

- **Parceria para Governo Aberto (Open Government Partnership - OGP):** el principal objetivo de OGP es difundir y fomentar a nivel mundial prácticas gubernamentales relacionadas con los principios de transparencia, participación social, responsabilidad e innovación.
- **Marco Civil da Internet (Ley N° 12.965/2014).**

Con respecto a la participación de la UFRGS en políticas de ciencia abierta y acceso abierto, en el 2018 destacan como temas priorizados del 4° Plan Nacional de Acción, establecer, de forma colaborativa, un modelo de referencia de política de datos abiertos que promoviera la integración, formación y sensibilización entre las sociedades y los tres ámbitos de gobierno desde la atención de las demandas sociales, y establecer mecanismos de gobernanza de datos científicos para el avance de la ciencia abierta en Brasil.

En cuanto a la Ley General de Protección de Datos Personales - LGPD (Ley N° 13. 709/2018), dispone sobre el tratamiento de datos personales, incluso en medios digitales, por una persona natural o por una persona jurídica de derecho público o privado, con el fin de proteger los derechos fundamentales de libertad e intimidad y el libre desarrollo de la personalidad.

Finalmente, para el año 2021, y con el 5° Plan de Acción Nacional de la Alianza para el Gobierno Abierto y con el proyecto REDEDADOSABERTOS, se invitaron a las Instituciones de Ciencia, Tecnología e Innovación (TIC) a colaborar para promover la Ciencia, Tecnología e Innovación, y con esto:

- Apoyar a los institutos brasileños de enseñanza e investigación en la apertura de datos de investigación.
- Fomentar el debate sobre la importancia de los repositorios de datos de investigación en las instituciones brasileñas.
- Planificar la estructuración de una Red Nacional Federada de Repositorios de Datos de Investigación, como la participación de los repositorios de datos del IBICT, CNPq, RNP e instituciones de investigación, entre ellas las cuatro TIC.

- Apoyar actividades de I+D y experimentación en repositorios de datos, investigar casos de estudio con conjuntos de datos de diferentes áreas de conocimiento.
- Apoyar acciones nacionales dirigidas a desarrollar una política de ciencia abierta.
- Ampliar las instituciones participantes del Consorcio Nacional para la Ciencia Abierta (ConCiencia) con la inclusión de las cuatro TIC, que podrán tener identificadores persistentes asociados a sus conjuntos de datos.

**MARGARITA LISOWSKA** (Universidad del Rosario, Colombia):

La Universidad del Rosario cuenta con 15 años de experiencia en la promoción de las iniciativas abiertas. Algunos de los avances que se han generado desde el 2018 hasta la fecha son:

- 2008: Repositorio Institucional
- 2010: Portal de revistas
- 2011: BIREDIAL-ISTEC
- 2016: Política institucional AA
- 2017: Firma de declaración de Berlín
- 2019: Ciencia Abierta

A nivel de políticas destacan para la gestión de datos de investigación en la UR Decreto Rectoral No. 1576 del 27 de marzo de 2019, que visibiliza la importancia que tiene para la institución la gestión de datos de investigación, y el compromiso de brindar un apoyo integral (recursos y servicios) a las personas investigadoras y lo que espera de los ellos como protagonistas del proceso.

En el tema de gestión de datos, la Universidad del Rosario cuenta con una Política Institucional para la gestión de datos de investigación, el repositorio de datos de investigación UR y servicios de apoyo a la gestión de datos de investigación.

Dentro de los servicios de gestión están distribuidos en un plan de gestión en donde se describe el proceso de manejo de datos durante y después de un proyecto de investigación. Y en el tema del repositorio de datos, se brinda acompañamiento durante el proceso de almacenamiento, y se capacita a las personas investigadoras para el uso adecuado y efectivo del mismo.

**MARISA DE GIUSTI** (Universidad Nacional de La Plata, Argentina):

La Universidad de La Plata inicia su trayectoria desde 1990, a partir de la asociación a ISTEC y firma en 1999 en la reunión de Santa Fe, Nuevo México, Estados Unidos como parte de ISTEC del lanzamiento de la OAI.

A partir de ese momento inicia una larga trayectoria, de la cual a nivel de legislación destaca la Ley N° 26.899 (2013) y su reglamentación, que tiene como objetivo el garantizar que la producción científico-tecnológica resultante del trabajo, formación y/o proyectos financiados total o parcialmente con fondos públicos y que haya sido publicada, aceptada para su publicación y/o atravesado un proceso de aprobación por una autoridad competente, se encuentre disponible en acceso abierto a través de repositorios digitales institucionales de acceso abierto, propios o compartidos. En sus artículos destaca la obligatoriedad de los organismos e instituciones públicas que componen el Sistema Nacional de Ciencia Tecnología e Innovación (SNCTI) de desarrollar repositorios digitales institucionales de acceso abierto.

Además, menciona el tema de la generación de datos primarios, documentos y/o publicaciones, considerando que dentro de sus cláusulas contractuales se debe presentar un plan de gestión acorde a las especificidades propias del área disciplinar, en el caso de datos primarios, y en todos los casos, un plan para garantizar la disponibilidad pública de los resultados esperados. Entendiéndose como dato primario a todo dato en bruto sobre los que se basa cualquier investigación y que puede o no ser publicado cuando se comunica un avance científico pero que son los que fundamentan un nuevo conocimiento.

Por otra parte, la legislación también abarca el tema de los repositorios digitales institucionales, los cuales deben ser compatibles con las normas de interoperabilidad adoptadas internacionalmente, y garantizar el libre acceso a sus documentos y datos a través de internet u otras tecnologías de información que resulten adecuadas a los efectos, facilitando las condiciones necesarias para la protección de los derechos de la institución y del autor sobre la producción científico-tecnológica.

En términos de las distintas partes que participan en los procesos, se han establecido una serie de roles. Se detalla a continuación en el caso de Argentina, como se han distribuido tales obligaciones.

El Ministerio de Ciencia, Tecnología e Innovación (Mincyt) es la autoridad de aplicación, entre sus principales funciones están la de establecer diferentes roles, apoyar a las instituciones y organismos, dictar directrices y velar porque se cumpla la ley.

Por su parte, los organismos e instituciones que componen el Sistema Nacional de Ciencia y Tecnología (SNCTI), se encargan de crear los repositorios de acuerdo a las directrices y sus políticas de acceso abierto.

En tanto que las agencias gubernamentales y de Organismos Nacionales de Ciencia y Tecnología del SNCTI son los encargados de establecer cláusulas de acceso abierto como contraparte de los financiamientos que otorga.

Finalmente, y en lo referente a los investigadores, tecnólogos, docentes, becarios de posdoctorado y estudiantes de maestría y doctorado, son quienes deben entregar o depositar su producción científica tecnológica al repositorio para difundirla en acceso abierto.

[Ver mesa de debate en YouTube](#)

# Mesa de debate: Evaluación abierta en revistas

## Una mesa para discutir sobre las posibilidades de la evaluación abierta en revistas científicas

MESA DE DEBATE

7 DE OCTUBRE DE 2022

### **RAFAEL GRANDE**

Revista *Encrucijadas. Revista Crítica de Ciencias Sociales* (España)

### **RIVA QUIROGA**

Revista *The Programming Historian* (Reino Unido)

### **ARTURO SILVA RODRÍGUEZ**

Revista *Digital Internacional de Psicología y Ciencia Social* (México)

### **MELISSA VARELA BRICEÑO**

Revista *Pensar en Movimiento* (Costa Rica)

### **ANDREA MÉNDEZ**

Revista *Arte, Cultura y Sociedad* (Costa Rica)

### **RESUMEN EXTENDIDO**

La discusión parte de los siguientes interrogantes:

*¿Cómo la evaluación abierta por pares se ejecuta en su publicación?*

*¿Cuál es la dinámica en el cual se enmarca este proceso de revisión abierta por pares?*

**RAFAEL GRANDE** (España):

En el caso de la revista *Crítica de Ciencias Sociales* de España, hay tres grandes ramas cuando se plantea una evaluación abierta:

1. La primera es la identidad abierta, quiere decir que haya un conocimiento entre la autoría y la persona que evalúa el artículo, es decir, un conocimiento de identidades.
2. Un modelo de informes abiertos, en donde los informes de evaluación estén disponibles posteriormente a la publicación para toda persona que lea el artículo.
3. Participación abierta, para que haya un sistema en la que una persona pone un artículo en abierto en una plataforma y toda la comunidad científica pueda solicitar evaluarlo, leerlo y hacer comentarios.

En el caso de la revista, actualmente está centrada solo en la identidad abierta, y esto se realiza a través del siguiente proceso. A partir de la preservación de artículos, el equipo editorial y los coordinadores realizan una preselección. Posteriormente, hay una evaluación abierta, donde se seleccionan evaluadores externos y a quienes se les hace llegar la propuesta de evaluar el artículo, conociendo el resumen y los autores. Posterior a esto, si la persona evaluadora acepta se le hace llegar el *paper* completo y se le notifica que su realimentación se le hará llegar a los autores del artículo con la identidad de quien lo ha evaluado. En este proceso participan, en algunos casos dos evaluadores, en casos excepcionales cuatro. En síntesis, se considera que el proceso ha permitido una realimentación interesante, sin que sea un punto en contra que se conozca la identidad de los evaluadores.

**RIVA QUIROGA** (Reino Unido):

El procedimiento es el siguiente: se pone a disposición un formulario, en el cual la gente envía una propuesta, la herramienta que se va a utilizar y la metodología. Posterior a esto, hay un primer diálogo con la jefatura de redacción, la cual, dependiendo del idioma, puede variar. Entonces, la persona

envía la propuesta, se hacen los ajustes, y una vez que la propuesta incluye toda la información relevante, pasa a un comité de revisión editorial que hace llegar sus recomendaciones por medio de correo electrónico.

La manera en que se refleja «lo abierto» ocurre en el uso de plataformas como GitHub, en la que se comparte código, y existe además un repositorio donde se trabajan las propuestas, y en el momento que se acepta una propuesta y se asigna la persona editora, se crea una sección que ofrece la posibilidad de hacer un foro de discusión que integra también los enlaces para la versión preliminar. Con esto se ofrece la posibilidad de generar una discusión en vivo. La persona editora hace una primera revisión en la cual deja el texto listo en el fondo para que quienes vayan a revisar, consideren solo los elementos centrales. Todos los comentarios se publican en abierto y luego participan los revisores. Tales pasos aseguran que exista un registro de todo el proceso.

**ARTURO SILVA RODRÍGUEZ (México):**

En el caso de México, se ha hecho un recorrido por los tres ámbitos: identidad abierta, informes abiertos, participación abierta.

Se trabaja con una aplicación que es el OJS, sin embargo, se cancelaron los comentarios abiertos debido a que no se disponía de un filtro adecuado. Lo que sí se mantuvo fue la identidad abierta de los evaluadores, así mismo, el informe que ofrecen las personas evaluadoras, este último se publica como un anexo de los artículos. De esta forma, la gente puede revisar el artículo y luego ver la valoración realizada.

Estos informes se construyen usando, por un lado, el índice de concordancia, que determina la confiabilidad entre evaluadores, la precisión en los juicios, es decir, se utilizan índices cuantitativos. Y lo que corresponde al análisis cualitativo, se hace a través de los comentarios que hacen los evaluadores en función de los rubros establecidos en la revista. En síntesis, no solamente se conoce la identidad de los evaluados y los autores de los artículos, sino que

también se hace público el informe de las personas evaluadoras y la conversación entre autores y evaluadores.

**MELISSA VARELA BRICEÑO** (Costa Rica):

En la revista se realiza la evaluación tradicional conocida como «a doble ciego» y la revisión abierta, en términos de conocer la identidad de las partes. De momento, en la revista no se publican los informes de revisión. Por otra parte, se trabaja con consentimiento tanto de autores como de evaluadores, cuando los autores envían su manuscrito junto con la carta de originalidad, comunican si están dispuestos a ofrecer su manuscrito para una revisión abierta. A partir de que el editor asociado inicia este proceso de revisión, a la persona evaluadora se le envía el resumen, el título y las palabras clave que le permitan identificar el grado de pertinencia, así mismo, se le solicita su consentimiento para realizar una evaluación abierta.

En cuanto a sistema de gestión, se utiliza OJS, el cual permite un grado de interacción, donde se conocen sus identidades y a la vez interactúen por medio del sistema para las observaciones pertinentes. También existe la posibilidad de que se genere una revisión mixta, cuando se realiza para un mismo documento una revisión abierta y otra a doble ciego. En ambos casos se les da crédito a los evaluadores, colocando su nombre en la portada del escrito y del editor asociado.

**ANDREA MÉNDEZ** (Costa Rica):

Para efectos de evaluación, la revista utiliza desde el inicio un sistema de arbitraje que es la evaluación abierta, en el que los pasos son: la recepción de las publicaciones por parte de la directora de la revista, posterior a lo cual, se hace una revisión ortotipográfica tras la cual se publica el preprint.

Además, la revista posee dos formatos para atender los preprint, y así mismo, las evaluaciones. Cuando el texto es de una persona externa a la

universidad, publicamos los preprints en el repositorio internacional Zenodo, mientras que si la persona pertenece a la universidad se publica el preprint en el repositorio institucional, lo mismo ocurre con las evaluaciones. Además, de abrir las identidades, y se sigue el procedimiento de repositorios, ya sea internacional o institucional.

[Ver mesa de debate en YouTube](#)

BIREDIAL  
ISTEC



---

Ponencias

---

Eje temático:

Comunicación académica, científica y cultural en abierto

[Ver bloque de ponencias en You Tube \(1\)](#)

[Ver bloque de ponencias en You Tube \(2\)](#)

# La innovación en la enseñanza y el aprendizaje de las artes visuales y plásticas en estudiantes de la EPEA-UJED

## **LETICIA ONTIVEROS MORENO**

Universidad Juárez del Estado de Durango (México)

[leonmoreno5@yahoo.com.mx](mailto:leonmoreno5@yahoo.com.mx)

## **JOSÉ CANDELARIO VÁZQUEZ MORENO**

Universidad Juárez del Estado de Durango (México)

[pincelcreativo@yahoo.com.mx](mailto:pincelcreativo@yahoo.com.mx)

## **SILVIA ISABEL NÁJERA TEJADA**

Universidad Juárez del Estado de Durango (México)

[najera.silvia@ujed.mx](mailto:najera.silvia@ujed.mx)

## **GABRIELA VITAL RUMEBE**

Universidad Juárez del Estado de Durango (México)

[vital.rumebe@ujed.mx](mailto:vital.rumebe@ujed.mx)

## **CLAUDIA GABRIELA GUERRA ROJAS**

Universidad Juárez del Estado de Durango (México)

[gabriela.querra@ujed.mx](mailto:gabriela.querra@ujed.mx)

## **RESUMEN**

La Escuela de Pintura Escultura y Artesanías (EPEA), de la Universidad Juárez del Estado de Durango (UJED), es una institución fundada hace más de sesenta años, el avance ha sido mayor en estos tiempos; por ello se plantea una comunicación en abierto de las actividades que se pueden aplicar para el desarrollo del proceso creativo innovador, el cual

da lugar a expresar formas de enseñanza y aprendizaje, que permite obtener mejores resultados, y por ende, es necesario aplicar habilidades diferentes en la práctica docente. El objetivo es describir las estrategias innovadoras, que fortalezcan acciones de enseñanza y aprendizaje, e impacten en la producción de obras creativas. El análisis de varios autores enfoca el proceso creativo y la innovación de la práctica docente, entre ellos: Marín, Rommel, Sambrano, Rodríguez, Cadme, que hacen mención a las técnicas innovadoras de las artes visuales y Rudolf Arnheim, que da la pauta a la enseñanza de las artes; se toman en cuenta además los principios de la UNESCO y lineamientos de la OCDE. La metodología aplicada es de tipo cualitativo, con un enfoque exploratorio, mediante el diseño de investigación documental. Los resultados obtenidos han sido favorables para el desarrollo del proceso creativo en estudiantes de las Artes Visuales.

#### **PALABRAS CLAVE**

Artes visuales, proceso creativo, innovación.

*Visual arts, creative process, innovation.*

## **Introducción**

La EPEA de la UJED surge a manera de una necesidad en la percepción del maestro Francisco Montoya de la Cruz, así como consecuencia de las misiones culturales derivadas de la propuesta de José Vasconcelos, y el origen en la Escuela Mexicana de Pintura en los años cuarenta.

Después de sesenta años de su fundación, se hace un recuento y se percibe como una necesidad la investigación en las Artes Visuales, a fin de generar conocimiento que permita dar un avance en la percepción del quehacer docente en el proceso de enseñanza y aprendizaje.

En este sentido se sabe que el arte involucra ideas que generan el proceso creativo y en esto intervienen en la creación de una obra de arte, donde se expresan emociones, sentimientos, estados de ánimo, se comunica y, por supuesto, se genera una experiencia estética que se comunica al espectador.

Son varias las organizaciones que hacen aportes a la innovación en la educación, como es la Organización para la Cooperación y el Desarrollo Económicos (OCDE), donde diferentes personalidades se reunieron en la Junta de Consejo Ministerial de la OCDE en 2010, y concluyen que la innovación y el conocimiento se están convirtiendo en factores claves para el crecimiento y bienestar; fueron personalidades de 38 países quienes analizaron los desafíos que implica dar a los estudiantes las habilidades para que alcancen un nivel de vida aceptable y generar motivación en el siglo XXI (WILNER et al., 2013).

Conscientes de que es necesario poner en práctica nuevos métodos de enseñanza y aprendizaje, en los que el estudiante no solo domine un cuerpo de conocimientos estéticos, sino que desarrolle las habilidades que le permitan ser creativo e innovador en la producción de obra plástica, nuestro trabajo pretende identificar las practicas didácticas deseables e innovadoras, para la producción artística de calidad, así como los diferentes procesos creativos aplicados a la experiencia estética.

El enfoque innovador y original de nuestra investigación consiste en la búsqueda y detección de estrategias didácticas que el docente aplique y que pueda ser útiles para el desarrollo del proceso creativo y, de esta forma, nuestros estudiantes obtengan una producción de obra de calidad.

El proceso creativo es un generador de ideas, las cuales son percibidas tanto por el artista al realizar una obra de arte, como por el espectador critico que se comunica por dicha obra; por lo tanto, conocer cómo se da ese proceso creativo en el alumno de las artes visuales de la EPEA-UJED, permitirá a los docentes formular nuevas estrategias de enseñanza para que el aprendizaje, como desarrollo creativo y crecimiento personal sea significativo y de mejor calidad.

La necesidad de promover y propiciar el crecimiento y desarrollo de los estudiantes en su individualidad, implica descubrir, vivir y motivar los valores inherentes a los seres humanos en un trabajo colegiado, pero para ello se necesitan estrategias didácticas diferentes e innovadoras que realmente desarrollen el proceso creativo del estudiante de las artes visuales y plásticas de la EPEA-UJED.

Conocer los cambios necesarios dentro del contexto de la educación artística, estableciendo modelos educativos centrados en las necesidades del estudiante, lo que permitirá generar estudiantes que apliquen y realicen innovación en su creación artística.

La necesidad de implementar nuevas y diferentes metodologías en un tiempo razonable, que permita una adecuada planeación, pero que sea lo suficientemente rápida para no quedarse atrás con respecto al resto del mundo.

Todo el cuestionamiento lleva a pensar en el desarrollo y percepción del proceso creativo que realizan los estudiantes de la carrera de Licenciatura en Artes Visuales y Plásticas; asimismo, es necesario conocer los contextos que favorecen e impulsan al desarrollo del proceso creativo y las estrategias innovadoras didácticas aplicables en la enseñanza y el aprendizaje que favorecen y fomentan el impacto en el crecimiento personal.

El objetivo es describir los procesos creativos que se han aplicado en la enseñanza de las artes visuales y plásticas en otras instituciones, así como conocer y analizar los contextos donde se fomenta y fortalece el proceso creativo en la producción artística para desarrollar estrategias innovadoras de enseñanza y aprendizaje y, de esta forma, realizar acciones que permitan la búsqueda de procesos didácticos y lograr un impacto en la producción de obra creativa de los estudiantes de la EPEA-UJED.

Por lo anterior, es de nuestro interés formular diferentes nuevas estrategias de enseñanza para el aprendizaje de las artes visuales y plásticas; buscar alternativas innovadoras en las prácticas docentes y garantizar que la

producción de los estudiantes tenga características de originalidad y se exprese en ellas el desarrollo creativo y personal que les permita la participación en exposiciones nacionales e internacionales, compitiendo en una sociedad cambiante y demandante de nuevos creadores de artes visuales y plásticas.

## Marco teórico

La investigación de la educación artística se da en diferentes vertientes: algunos estudiosos del tema se refieren a un espacio heterogéneo y con límites hasta cierto punto difusos; por ejemplo, en América Latina, la investigación de la educación artística comprende la enseñanza de todas las disciplinas artísticas: la música, la danza, el teatro, las artes visuales, entre otras, mientras que en Europa y Estados Unidos de Norteamérica, el mismo término únicamente hace referencia a la enseñanza de las artes visuales. Y es así que algunos investigadores se dedican a la educación artística junto con otras actividades como es la docencia, la creación artística, la psicología, la sociología o la crítica de arte (MARÍN, 2011).

En este sentido, se mencionan algunos teóricos que realizan investigación de la educación de las artes y cuyos aportes son la base para el planteamiento de nuevas estrategias de enseñanza para el aprendizaje.

Los planteamientos de organizaciones internacionales como es la UNESCO, son un especial aporte al desarrollo de nuevas estrategias de enseñanza y aprendizaje de las artes:

Las sociedades del siglo XXI exigen cada vez más trabajadores creativos, flexibles, adaptables e innovadores y los sistemas educativos deben ajustarse a esta nueva situación. La educación artística constituye asimismo un medio para que los países puedan desarrollar los recursos humanos necesarios para explotar su valioso capital cultural. La utilización de estos recursos y este capital es vital para los países si desean desarrollar industrias e iniciativas culturales fuertes, creativas y sostenibles, las cuales pueden desempeñar un papel clave al

potenciar el desarrollo socioeconómico en los países menos desarrollados (ORGANIZACIÓN DE LAS NACIONES UNIDAS PARA LA EDUCACIÓN, LA CIENCIA Y LA CULTURA, [UNESCO], 2006).

Varios autores han realizado investigaciones en relación a las aportaciones de Elliot Eisner, entre ellos, Juanola y Masgrau, quienes mencionan que, bajo una actitud abierta, autocrítica y transformable, se debe continuar la corriente teórica de Eisner, y aplicar su legado: “aprender a vivir el mundo con un marco de referencia visual, con una plena perspectiva estética y defender en todo momento la importancia del arte en la vida, en la educación y en la investigación” (JUANOLA & MASGRAU, 2014).

En el 2018, Cadme menciona que las técnicas innovadoras de las artes visuales se establecen mediante el uso de remix digital, que consiste en una mezcla de producciones sonoras, textuales, visuales o audiovisuales creadas por otras personas. Actualmente es una práctica habitual en la red, de donde se toman prestadas producciones preexistentes para crear otras nuevas a partir de un proceso de “remezcla”. *Stop motion* es una técnica de animación que consiste en capturar y reproducir una serie de imágenes fijas sucesivas, lo que da como resultado la impresión de movimiento de los objetos estáticos. Videoarte, expresión artística formada a base de imágenes en movimiento y audio. Puede adoptar distintas modalidades: grabaciones que se distribuyen en la red, en cintas, en videos, etc.; videos exhibidos en galerías u otros espacios expositivos; instalaciones escultóricas, que pueden incorporar uno o más aparatos de televisión o monitores de video; o actuaciones en las que se incluyen proyecciones de video (CADME, 2018).

Las estrategias didácticas para estimular el proceso creativo en los estudiantes deben estar en base a que los profesores apliquen en su práctica docente el uso de herramientas digitales y materiales, pero como si fuera una experiencia pictórica, donde están implícitas las emociones y los sentimientos, no solo una didáctica aplicada a la enseñanza de las artes visuales y plásticas sino a una enseñanza de valores, de despertares de conciencia de sentir el

alma, el ser y la propia esencia de los que se enseña, para poder desarrollar procesos creativos en los estudiantes.

Rudolf Arnheim es un teórico que ha aportado varias investigaciones al pensamiento visual y la educación artística, y menciona que:

Los profesores de arte caen fácilmente en la tentación de enseñar demasiado pronto a sus estudiantes trucos como la perspectiva central, al profesor le hace parecer profesional, al estudiante que quiere emular los patrones adultos le agrada y a los padres les impresiona; pero con demasiada frecuencia esto no es más que un torpe lavado de cerebro; se ignora el hecho de que la perspectiva central, es una fórmula muy especial que surgió por primera vez en el siglo XV, cuando la búsqueda intuitiva de una nueva centralización de las formas en la dimensión de profundidad había alcanzado un estadio en el que esta técnica geométrica apareció como la consecuencia lógica final (ARNHEIM, 1993).

En esta propuesta de investigación, esperamos no caer en enseñar la perspectiva central que se menciona en el párrafo anterior.

Parafraseando a Arnheim, el mejor profesor de artes será aquel que con la sabiduría de un buen jardinero, observa, juzga y echa una mano cuando su ayuda es necesaria, no aquel que se guarda todo para sí mismo, que no comparte lo que sabe (ARNHEIM, 1993).

No queda otra más que fomentar, difundir y aplicar estrategias innovadoras para la enseñanza de las artes visuales y plásticas en la EPEA-UJED.

Hay estrategias de enfoque, expresión y creación, donde el componente principal tiene relación directa con la noción de creatividad; este modelo recoge los conceptos ligados a la expresión personal creativa (RAQUIMÁN & ZAMORANO, 2017).

Un aspecto importante y común de estos enfoques es que la estrategia pedagógica exige que el estudiante se conecte con su sensibilidad (esta cambia continuamente), la cual genera la expresión de impulsos personales estimulando el proceso de la producción creativa, prescindiendo de directrices

que se desprenden de la formación tradicional o de estructuras formales de la enseñanza de las artes visuales y plásticas.

La asistencia museica para la aplicación del aprendizaje a través de la pedagogía es un programa propuesto por el Museo Guggenheim en Estados Unidos y España, en el cual, mediante la participación de artistas visuales y el desarrollo del proceso creativo, los alumnos adquieren competencias de planificación y gestión de proyectos para los que trabajan en equipo y se ven obligados a utilizar el pensamiento crítico; este programa es un modo de llegar a niños que quizá tendrían problemas para seguir los métodos de enseñanza tradicionales y, lo que es más importante, en general estimulan la autoestima y el crecimiento personal de los niños (ORGANIZACIÓN DE LAS NACIONES UNIDAS PARA LA EDUCACIÓN, LA CIENCIA Y LA CULTURA, [UNESCO], 2006).

En México, se encontró que la Escuela Nacional de Pintura, Escultura y Grabado "La Esmeralda" tiene como visión ser una institución educativa de vanguardia, que, a través de la intersección propositiva entre las disciplinas tradicionales, los medios tecnológicos de innovación y sus prácticas creativas, generen redes con otras instituciones y comunidades, donde los estudiantes sean capaces de construir con conciencia y capacidad crítica (INSTITUTO NACIONAL DE BELLAS ARTES Y LITERATURA, 2021).

ROMMEL, en el 2014, señala que se deben cambiar las escuelas de artes y acercarlas a la realidad de afuera, un hecho doloroso es comprobar la dificultad que los egresados de artes tienen para hacer su trabajo y vivir de él; hay que hacer muchos muchísimos cambios en la educación superior, paralelos a la realidad tecnológica, económica y social, pero no solo en proyectos de papel ni como intenciones aisladas sino con una función indagadora, creadora y divulgadora de sentidos (ROMMEL *et al.*, 2014).

La Universidad Nacional Autónoma de México (UNAM), en su página web menciona que el estudiante deberá poseer una afinidad por la investigación-producción, elemento fundamental para su actividad misma, así como promover la innovación en cuanto a lenguajes, técnicas y materiales, y coordinar y fundamentar proyectos que sean una aportación de expresión

artística que esté orientada al impacto en el desarrollo humano, artístico y cultural (UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO, 2021).

La EPEA-UJED menciona en su Plan de estudios que el estudiante desarrolla innovaciones y propone soluciones a problemas a partir de métodos establecidos (NÁJERA, 2011). En la visión que se tiene de la proyección nacional e internacional dirigida al perfeccionamiento y actualización en la enseñanza de las artes visuales, sustentándose en el talento artístico y artesanal para buscar los más altos valores estéticos, calidad y excelencia (UNIVERSIDAD JUÁREZ DEL ESTADO DE DURANGO, 2021). Y el objetivo es formar personas capaces de desempeñarse en la disciplina artística de la pintura, la escultura y el grabado, buscando una capacitación que les permita desarrollarse en estos ámbitos de una manera exitosa en los aspectos profesionales, logrando un mejor desarrollo cultural y social.

## Metodología

Esta investigación es cualitativa, con un enfoque exploratorio, y un diseño de investigación tipo documental, donde las metodologías aplicadas en la investigación de las artes visuales y plásticas son hasta cierto punto nuevas, diferentes e innovadoras, por las formas de diseñar protocolos de investigación en educación artística. La idea básica de toda metodología es aproximar los usos y tradiciones profesionales de la creación artística a las normas y criterios de la investigación en ciencias humanas y sociales, de tal forma que en lugar de considerar la actividad científica como contradictoria y opuesta a la actividad artística, lleguen a considerarse complementarias, ya que cada una ofrece enfoques y resultados diferentes sobre los problemas humanos y equivalentes, disfrutando de vsus peculiares técnicas y procedimientos de indagación así como sus logros cognoscitivos de un reconocimiento académico semejante. En todos los casos, se trata de usar el gran bagaje de conocimientos y estrategias profesionales de análisis, representación y persuasión, que son característicos de las creaciones artísticas, para abordar los problemas educativos.

Las metodologías artísticas de investigación plantean muchos problemas epistemológicos y metodológicos, de los que se destacan tres:

**a)** La pluralidad de lenguajes de representación y presentación de los datos, procesos y resultados, que van más allá del lenguaje verbal escrito y de los análisis estadísticos de datos cuantitativos.

**b)** La complejidad semántica y la amplitud connotativa de los resultados son distintivas de los lenguajes artísticos; esta apertura de significados de las obras artísticas es interpretada en ocasiones como ambigüedad o ausencia de significados y por tanto impropia para la investigación.

**c)** La flexibilidad con los datos empíricos; en este sentido, las metodologías cuantitativas dependen de los datos exactos que se obtienen de test o de cuestionarios. Sabemos que la observación participante o las entrevistas a profundidad proporcionan datos más dúctiles y maleables porque dependen tanto de los sucesos y de la idiosincrasia de las personas entrevistadas como de las capacidades de empatía y dialogo de la persona que hace la entrevista (MARÍN, 2011).

## Resultados

Las estrategias didácticas para el desarrollo de las artes visuales y plásticas dependerán directamente de la práctica docente que cada profesor tenga y aplique en el aula presencial, virtual o híbrida; definitivamente es de considerar que la enseñanza de las artes tiene un sinfín de estrategias innovadoras, las cuales puedes ir surgiendo en la clase, por ejemplo:

- El uso de herramientas digitales.
- Escuchar música del agrado de los estudiantes mientras se desarrolla la práctica docente.
- Realizar actividades didácticas para el desarrollo del proceso creativo es sin duda alguna las estrategias que se plantean en la siguiente propuesta.

Se llegó al planteamiento de un programa para el desarrollo del proceso creativo, donde se plantean el cómo, el qué y el para qué. Las actividades planteadas fueron veinte; por cuestiones de espacio solamente se describen cinco: se aplicaron a varios grupos obteniendo excelentes resultados, como el desbloqueo de los procesos creativos de los estudiantes de artes visuales, donde ellos manifiestan que, al realizar algún ejercicio propuesto en esta REA, adquirieron y reconocieron la conceptualización de valores importantes como es la responsabilidad, la disciplina y el trabajo en equipo.

El cambio de entusiasmo y compromiso fueron respuestas inmediatas, que se observaron al presentar un trabajo académico o una obra plástica, los estudiantes mencionaron que fue un proceso dinámico y creativo que permitió el despertar, la capacidad de asombro que quizás tenían dormida en algunos casos y en otros se desconocía.

Otro aspecto importante de cambio fue la forma de realizar una obra plástica ya no mecánicamente sino con un propósito e integrando un objetivo que los guía y les va dando la pauta a continuar, sin llegar a la mecanicidad del proceso creativo.

Así mismo se observó que los estudiantes que participaron en estas dinámicas del REA, se volvieron más participativos en otras actividades, así como entablaron un compromiso de responsabilidad para con ellos mismos.

Las siguientes actividades son ejercicios prácticos para el desarrollo del proceso creativo:

#### **EJERCICIO 1: AUMENTO DE LA CREATIVIDAD**

Piense en una idea innovadora (un diseño, un objeto, un proyecto, un plan, un programa) y búsquele visibilidad. Concéntrese en imágenes visuales que pueden definirla. Créelas. Deje fluir su fantasía. Váyase a las nubes. Suspenda todo pensamiento lógico-racional. Cierre los ojos. Relájese. Respire rítmicamente... visualice sus tensiones saliendo de su cuerpo, fluyendo a través de su cuello, de los hombros de los brazos hacia sus manos y de su vientre, coxis y muslos hacia sus pies. Y finalmente saliendo de sus dedos al exterior. (RODRÍGUEZ, 1999, p. 168)

### EJERCICIO 2: TE REGALO ESTAS FRASES

Se hace una lectura a los estudiantes:

*Mi pensamiento está hecho de ideas originales, las notas musicales de mis palabras son la melodía de mi mente. Pienso, luego creo (de creer y de crear).*

*Soy el director de mi orquesta mental, mi voz interior es rica y poderosa. Tengo muchas palabras para expresar los matices de mis sentimientos. Yo elijo mis pensamientos, mis sentimientos y mis reacciones y tengo la capacidad de aceptarme tal como soy; necesito el sonido de la naturaleza y tengo la capacidad necesaria para dirigir mis pensamientos.*

*Los ojos de los niños son el espejo del corazón de la tierra, soy capaz de volar con el pensamiento, cada vez que puedo le regalo un libro a alguien a quien quiero, escribo mis sueños, mis recuerdos, mis vivencias. Soy capaz de expresar los sentimientos más profundos de mi corazón. El sol me entrega su energía, yo la administro con amor y con eficacia. Soy imaginativo, soñador y realizador. Soy capaz de visualizar mis metas con claridad; la imaginación es el poder activo de mi cerebro, la imaginación otorga fuerza a la realidad. Soy feliz.*

*El sabor del amor es tan dulce que puedo sentirlo en mi boca, soy el director de la orquesta de mi vida y la felicidad es un estado de conciencia que soy capaz de sentir. (SAMBRANO, 2000, p. 105)*

### EJERCICIO 3: VISUALIZACIÓN IMAGINATIVA. IMAGINAR MEJOR

1. Entrena tu imaginación para que imagines lo que desees y no te limites solo a recibir lo que venga, puedes adiestrarte en visualizar, para escuchar y sentir lo que necesite.
2. Trazas un objetivo, lo imaginas y luego actúas.
3. Cuando comiences a imaginar dale más brillo, precisión y permite que las imágenes sean cada vez más nítidas tanto como lo desees.
4. Presta atención a las imágenes que aparecen en el horizonte mental, deja que fluyan cada vez más nítidas.
5. Anota las imágenes que te hayan llamado la atención, las que te brindan inspiración y motivos para hacer que tu vida sea cada vez mejor.
6. Toma conciencia de alguna imagen mental que aparezca y no quieras tener presente.
7. Cambia las imágenes que no te satisfacen; eres como un cineasta, filma las proyecciones que crea tu mente. (SAMBRANO, 2000, p. 49)

#### EJERCICIO 4: EXTERNAR EMOCIONES

Se proyectan 12 diapositivas de "rostros" en secuencias de seis o bien en fotografías. Después de cada bloque se les deja un par de minutos para que escriban sus reacciones. Al final todos reportan sus vivencias al respecto.

Se realiza un ejercicio análogo con diapositivas donde se proyecten imágenes de «manos»

Identifíquese con:

Un color (si yo fuera un color qué color sería); un animal (si yo fuera un animal); una flor (si yo fuera); un libro (si yo fuera); un juguete (si yo fuera); una estación del año (cuál); un país (cuál).

Explique el porqué de sus respuestas en cada caso; ayúdese de estas respuestas para perfilar y definir su autoimagen. (RODRÍGUEZ, 1999, p. 81)

#### EJERCICIO 5: IMAGINANDO PERSONAS

Visualiza a todas las personas con las que has conversado hoy: ¿cómo llegan a tu mente? ¿están sentadas? ¿caminando? ¿de qué color son sus ojos, su pelo? ¿cuál es su tamaño? ¿qué edades tienen? ¿cómo están vestidas?

Imagina alguna característica importante de alguna de ellas.

Visualiza a todas las personas con las que estuviste ayer, la semana pasada, el mes pasado, el año pasado.

Visualiza a tus compañeros de clase, de otros grupos.

Piensa en tus compañeros de trabajo (si trabajas).

Visualiza a tus hijos (si eres casada o casado).

Imagina a alguien que quieres conocer.

Imagina el físico de alguien que solo conoces por teléfono (voz), cómo visualizas sus manos, sus ojos, sus pies.

Visualiza a todas las personas que han pasado significativamente en tu vida, deja que desfilen en tu imaginación. (SAMBRANO, 2000, p. 52)

## Conclusiones

Definitivamente, la búsqueda de estrategias didácticas innovadoras para lograr el desarrollo del proceso creativo, ha dejado inquietudes y necesidades, así como alternativas, para continuar investigando sobre todo en lo que respecta a los ejercicios que se pueden aplicar para el logro de los objetivos propuestos; sin embargo, sabemos que no es el parteaguas a continuar, ya se ha mencionado en otras ocasiones pero sin darle la importancia a realizar las

diferentes alternativas que se presentan para que realmente se logre el desarrollo del proceso creativo en estudiantes de las artes visuales y plásticas de la EPEA-UJED.

**LETICIA ONTIVEROS MORENO**

El proceso creativo interviene como generador de ideas, las cuales son percibidas por el artista para poder innovar en la realización de una obra de arte. Para fomentar la creatividad en el estudiante, es necesario conocer el contexto artístico, dentro del que se desarrolla, sus necesidades y las aplicaciones prácticas del medio en su desarrollo, lo cual nos permitirá realizar una planeación innovadora de los contenidos formativos que habrá de desarrollarse. En este sentido, se realizó una investigación cualitativa, buscando describir los procesos creativos, que se han aplicado en la enseñanza de las artes plásticas y visuales en otras instituciones, mencionar los contextos formativos de nuestros estudiantes, y realizar acciones que permitan la búsqueda de estrategias innovadoras en la enseñanza que fomenten el desarrollo de la creatividad. Encontramos que todo proceso educativo debe servir como generador de ideas en los educandos, que estimule su capacidad creativa, y su capacidad de involucrarse con el espectador. Sin embargo, este proceso debe ser motivador, para que el alumno sea quien descubra sus capacidades e invente el cambio aplicando los recursos del medio ambiente, a través de un conocimiento práctico no sólo de las técnicas y métodos artísticos, sino del sentimiento del espectador y de cómo poder motivar con su obra.

**GABRIELA VITAL RUMEBE**

La innovación en la enseñanza es una necesidad fundamental y debe ser de interés de toda institución formadora de profesionales en las artes el incrementarla, esto a partir del desarrollo de las habilidades en el proceso creativo, dándole la importancia como actividad esencial. Este estudio se

centra en su indagación y se concluye que el trabajo parte de la individualidad del artista y que la motivación es diferente en cada persona, por lo tanto, las acciones tendrán que ser encaminadas a implementar actividades adaptables a las distintas formas de aprender y de crear, de igual forma los contextos son esenciales en la búsqueda de espacios creativos, así como la implementación de las nuevas tecnologías para el desarrollo de las habilidades. Por lo anterior se propone un plan de acción, que contiene estrategias nuevas, para que sean implementadas en las escuelas y como mencionan los teóricos en los que se basó este estudio educar en la expresión personal, a partir de un pensamiento crítico generando conciencia social que impacte en el desarrollo de la sociedad.

**SILVIA ISABEL NÁJERA TEJADA**

En la actualidad, una de las preocupaciones de la EPEA- UJED en la LAVyP es implementar nuevas prácticas educativas. Los diferentes enfoques respecto a la enseñanza y desarrollo creativo de la percepción de los docentes permiten articular la construcción de marcos referentes de innovación en los estudiantes. Esta construcción de marcos basados en las diferentes prácticas precisa algunas características realistas, las cuales encaminan a la implementación de nuevas formas de construir la producción artística, para desarrollar la creatividad en el talento artístico y artesanal de sus alumnos para alcanzar los más altos valores de calidad.

**CLAUDIA GABRIELA GUERRA ROJAS**

El desarrollo del proceso creativo adquiere una mayor relevancia y como una necesidad del siglo XXI para hacer frente a las necesidades propias de una sociedad invadida por la tecnología. Durante el desempeño como docente por más de treinta años, se han implementado diferentes estrategias para el aprendizaje de las artes visuales, esto ha permitido observar intención en trabajar en dos vertientes: primero en la elaboración de una obra derivada de las vivencias y recuerdos de manera libre donde el alumno se expresa, disfruta

y siente cada pincelada de color y cada línea con el lápiz, en este sentido la recomendación ha sido “disfrutar el viaje” (el proceso de cada trazo, sin pensar en cómo va a quedar la obra); permitir que el estudiante fluya y en ocasiones no atienda a las reglas establecidas por el canon artístico, que se sienta satisfecho ha sido la clave para el desarrollo del proceso creativo. Y la segunda vertiente está enfocada trabajar para sí mismo y no para darle gusto a los demás (nunca terminaría de darle gusto a todo el mundo), esto es que el estudiante produzca su obra dándose gusto a sí mismo y vivir la experiencia.

**JOSÉ CANDELARIO VÁZQUEZ MORENO**

## **Bibliografía**

- ARNHEIM, R. (1993). Consideraciones sobre la educación artística, 3, (1ª), 56.
- BENNASSAR, LI. M. (1994). Desarrolle su creatividad, (1),81.
- BETANCOURT, M. (1996). Psicología y creatividad, apuntes y reflexiones, (1), 45.
- CADME, G. (2018). Estrategias didácticas en la formación docente a través de las artes visuales. Universidad Católica de Cuenca, España, IV Encuentro Internacional de Ciencias Pedagógicas de Ecuador.
- INSTITUTO NACIONAL DE BELLAS ARTES Y LITERATURA. (2021). Misión de la Esmeralda. <https://www.esmeralda.edu.mx/mision>
- JUANOLA T. Y MASGRAU J. (2014). Las aportaciones de E. W. Eisner a la educación: un profesor paradigmático como docente investigador y generador de políticas. *Revista Española de Pedagogía*, LXXII(259). <https://revistadepedagogia.org/wp-content/uploads/2014/09/259-06.pdf>
- MARÍN V. R. (2011). Las investigaciones en educación artística y las metodologías artísticas de investigación en educación. *Temas, Tendencias y Miradas*, 34, (3ª). <http://www.redalyc.org/articulo.oa?id=84820027003>
- NÁJERA, T. (2017). Plan de estudios de la Licenciatura en artes visuales y plásticas, EPEA-UJED.

- RAQUIMÁN O. P. Y ZAMORANO, M. (2017). Didáctica de las Artes Visuales, una aproximación desde sus enfoques de enseñanza. *Estudios Pedagógicos*, 43(1): 439-456.  
[https://www.scielo.cl/scielo.php?pid=S0718-07052017000100025&script=sci\\_arttext](https://www.scielo.cl/scielo.php?pid=S0718-07052017000100025&script=sci_arttext)
- RODRÍGUEZ, E. M. (1999). Pensamiento Creativo, *Integral*, (1), 4-168.
- ROMMEL, S., GUZMÁN, A. ET AL. (2014). *La educación superior en artes. Reflexiones, consideraciones y propuestas*. Facultad de Artes Plásticas de la Universidad Veracruzana.
- SAMBRANO, J. (2000). Cerebro: Manual de Uso, (1), 31-127.
- UNIVERSIDAD JUÁREZ DEL ESTADO DE DURANGO. (2021). Identidad institucional, EPEA.  
<https://www.ujed.mx/escuela-de-pintura-escultura-y-artesanias/identidad-institucional>
- UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO. (2021). Oferta Académica Artes Visuales.  
<http://oferta.unam.mx/artes-visuales.html>
- UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION [ORGANIZACIÓN DE LAS NACIONES UNIDAS PARA LA EDUCACIÓN, LA CIENCIA Y LA CULTURA] UNESCO (2006). Hoja de ruta para la educación artística. Conferencia mundial sobre educación artística: construir capacidades creativas para el siglo XXI Lisboa.  
[http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/CLT/pdf/Arts\\_Edu\\_RoadMap\\_es.pdf](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/CLT/pdf/Arts_Edu_RoadMap_es.pdf)
- WILNER, E.; GOLDSTEIN, T. R.; & LANCRIN, S. V. (2013). ¿El arte por el arte?, la influencia de la educación artística, OCDE.  
<https://www.ipn.mx/assets/files/innovacion/docs/libros/el-arte-por-el-arte/capitulo-1-la-influencia-de-la-educacion-artistica.pdf>

# Produção científica sobre os princípios FAIR a partir do Directory of Open Access Journal (DOAJ)

## **ANA JULIA LOPES**

Universidade Federal do Rio Grande do Sul (UFRGS) (Brasil)

[anajuliafurg@gmail.com](mailto:anajuliafurg@gmail.com)

## **EDNA KARINA DA SILVA LIRA**

Universidade Federal de Santa Catarina (UFSC) (Brasil)

[liraa.karina@gmail.com](mailto:liraa.karina@gmail.com)

## **ANGÉLICA CONCEIÇÃO DIAS MIRANDA**

Universidade Federal do Rio Grande (FURG) (Brasil)

[angelicacdm@gmail.com](mailto:angelicacdm@gmail.com)

## **RESUMEN**

Uma mobilização em prol da abertura dos dados científicos vem sendo discutida visto que novas tecnologias estão surgindo facilitando o acesso, transferência e análises de dados. Nesse contexto, surgem os princípios orientadores FAIR com o intuito de auxiliar no processo de gestão de dados de pesquisa. Diante do exposto o presente estudo teve como objetivo investigar a produção científica sobre os Princípios FAIR a partir do Directory of Open Access Journals - DOAJ. Utilizou-se da técnica bibliométrica para desenvolver os objetivos. A coleta de dados aconteceu entre os dias 26 a 30 de agosto de 2021. A busca recuperou 103 documentos - 97, com o termo em inglês, e 6 em português, sendo salvos no Zotero para então começar a análise. Constatou-se que, dos 103 documentos recuperados no DOAJ, 40

disseram sobre o tema da pesquisa. Entre os principais resultados se destaca que os periódicos Data Science Journal, Frontiers in Marine Science, International Journal of Digital Curation, Liinc em Revista e PeerJ Computer Science mais publicaram sobre o tema; os assuntos abordados dentro da temática foram Princípios FAIR; Dados FAIR; Repositório de Dados; Ciência Aberta; FAIR; Gestão de dados; interoperabilidade; Infraestrutura de pesquisa e Gerenciamento de dados de pesquisa. As considerações finais mostram que o DOAJ comporta amplamente os debates difundidos sobre o assunto pelas comunidades internacionais científicas nos últimos anos.

#### **PALABRAS CLAVE**

Princípios FAIR; gestão de dados de pesquisa; DOAJ.

## **Introdução**

A Ciência Aberta é um movimento que nos permite colaborar e compartilhar resultados de pesquisa à comunidade científica, ocorrendo desta forma a democratização do acesso à informação. O acesso livre ao conhecimento produzido é a forma mais tradicional dentro desse movimento, visto as inúmeras ferramentas, como repositórios, portais de periódicos, diretórios e outras redes de compartilhamentos que podem servir de exemplo e contribuíram para o crescimento da produção científica e uso de dados. (BIAZON; MARIN, 2016).

Uma das principais propostas da Ciência Aberta é a abertura dos dados científicos, uma vez que o crescimento da produção e do uso de dados nesse contexto têm proporcionado discussões acerca de sua importância, uso, desafios e impactos em uma sociedade conectada e amparada pelo usos das tecnologias digitais cada vez mais eficientes e especializadas, e a adoção de diferentes práticas, que visam os interesses sociais e coletivos, além do

compartilhamento e reutilização dos dados e informação, são proposições que colaboram para a descoberta de novos conhecimentos. (HENNING *et al.*, 2019b).

A tecnologia colaborou para maximizar o alcance da produção científica e do uso dos dados científicos. Com o debate sobre sua importância, surgiram então os Princípios Orientadores FAIR com o intuito de auxiliar no processo de gestão, compartilhamento e reuso dos dados de pesquisa. SANTOS, ALMEIDA e HENNING (2017) afirmam que, com os dados bem descritos, estruturados e disponíveis para utilização é possível aumentar a capacidade e descobertas de novas pesquisas.

Com o crescimento da produção científica e a (re) utilização dos dados de pesquisa ocorreu o impacto do acesso aberto na comunicação científica, ampliando o acesso ao que é produzido internacionalmente.

Sob esse viés, o foco deste estudo foi mapear a produção científica acerca dos Princípios FAIR utilizando o Directory of Open Access Journals -DOAJ, que devido ao seu alcance mundial contribui para a disseminação das discussões acerca das áreas do conhecimento, identificar as revistas que mais publicaram sobre o tema bem como sua localização geográfica e conhecer os assuntos mais abordados dentro da temática (palavras chaves).

A pesquisa teve como procedimento técnico a bibliometria e utilizou duas de suas três leis: a Lei de Bradford, na qual permite calcular o grau de relevância dos periódicos em uma determinada área do conhecimento, e a Lei de Zipf, que possibilita estimar as frequências de ocorrência das palavras de um determinado texto científico e tecnológico (RODRIGUES & VIEIRA, 2016).

## **Acesso Aberto e Ciência Aberta**

O compartilhamento de conhecimento tem sido o pilar da produção e inovação de novos conhecimentos. No campo científico a informação científica é cada vez mais importante para o avanço das ciências, de modo que limitar o acesso aos resultados de pesquisa pode ocasionar danos à sociedade

(GUMIEIRO & COSTA, 2012). As iniciativas do acesso aberto é um movimento que combate situações nesse cenário.

O contexto que levou a origem desse movimento foi a mudança de situações sociais e avanços tecnológicos, nas quais as mais significativas foram, de acordo com MEADOWS (1999), o surgimento dos periódicos científicos, das comunidades científicas, da qualificação da pesquisa bem como especialização nas áreas do conhecimento, além do surgimento da era eletrônica.

COSTA e LEITE (2016) corroboram que somados aos fatores supracitados surgem outros dois elementos para as transformações no sistema de comunicação científica: o primeiro é a insatisfação dos pesquisadores com a forma tradicional das publicações dos periódicos científicos, onde tinham que pagar para ter acesso aos resultados de pesquisas, o que instituiu barreiras para o acesso do que produziam. O segundo elemento é o surgimento de meios para acelerar e dinamizar os processos de comunicação, concedido por tecnologias mais avançadas, que já vinham sendo incluídos desde os séculos anteriores no universo da comunicação científica. A junção desses fatores foi o cenário para o surgimento do movimento de acesso aberto.

O acesso aberto se refere a disponibilidade gratuitamente por qualquer pessoa a resultados de pesquisa científica, baseando-se na ideia de que o conhecimento científico é público e por isso deve estar acessível para todos, sendo fundamental que a informação seja disponibilizada em plataformas digitais (AGÊNCIA..., 2022). BRANDÃO, MOREIRA e TANQUEIRO (2021) complementam que nesse se considera como informação científica artigos, fatos ou números recolhidos para investigação ou discussão, dados de investigação (tratados e não tratados), estatísticas, resultados de experiências, entre outros. Reiteram ainda que desde o início foi entendido que uma parte maior do conhecimento de forma aberta e transparente, no quesito da democratização, seria positivo à toda sociedade.

Com esse objetivo —de disponibilizar livre e integral os resultados das pesquisas científicas, em texto completo, por meio da internet— e com ideias que foram se fortalecendo a ideia do movimento de acesso aberto chamado hoje de ciência aberta.

Uma das iniciativas mais importantes para a consolidação desse movimento foi a Iniciativa de Acesso Aberto de Budapeste (BOAI - Budapest Open Access Initiative) que desencadeou uma campanha mundial em prol do acesso aberto a todas as publicações científicas revisadas por pares. Houve outras iniciativas e declarações em apoio, tais quais podemos citar a Declaração de Bethesda em 2003, que ampliou o conceito, orientações e recomendações às instituições, organizações, pesquisadores, bibliotecários e editores, e a Declaração de Berlim, também em 2003, que reforçou os critérios já firmados e destacou o uso da internet como principal ferramenta (QUEIROZ; ARAUJO; NASCIMENTO, 2019; AGÊNCIA..., 2022; COSTA, 2006).

Um novo desafio foi lançado para a comunidade científica, salienta SANTOS, ALMEIDA, HENNING (2017). De acordo com as autoras (2017, p. 11) esse desafio tem como destaque o aumento de questionamentos sobre o papel da ciência e sua verdadeira contribuição para o desenvolvimento de sociedades mais justas e sustentáveis, surgindo então uma “urgente necessidade de revisão e criação de um novo modelo de organização da ciência pautado numa perspectiva mais colaborativa e aberta, num movimento que se denomina Ciência Aberta”.

Tem-se aí novas frentes, como os dados científicos abertos, ferramentas científicas abertas, como software e hardware, os cadernos abertos de laboratório, a educação aberta e a ciência cidadã. Mais do que um termo ‘guarda-chuva’, a ciência aberta avança no sentido de integrar essas várias frentes, ampliando a questão do acesso à informação científica para focar também nas novas formas de produção, circulação e apropriação social da informação e do conhecimento em ciência, tecnologia e inovação (ALBAGLI, 2017).

O Manual de Formação em Ciência Aberta (2018) confirma que esse movimento é a prática científica que permite que outros possam colaborar e contribuir, de forma que os dados de investigação, e demais processos, são livremente disponibilizados com condições de proporcionar o reuso, reprodução e redistribuição da investigação e dos dados e métodos implícitos

Entende-se, então, que o Movimento da Ciência Aberta foca nos princípios de que a investigação científica deve ser partilhada e reutilizada com a comunidade científica e sociedade em geral, de forma que, se antes, com o Movimento de Acesso Aberto, era fundamental comunicar os principais resultados de pesquisa e disponibilizá-los livremente na internet, nesta ocasião torna-se importante registrar toda a etapa da pesquisa e torná-la acessível.

## Princípios FAIR

O crescente anseio pelos dados gerados pelas atividades científicas proporcionou, tanto por agências de fomento, instituições de pesquisa e comunidade acadêmica, a necessidade de uma gestão ativa superior ao simples armazenamento e disponibilização na web, um sistema confiável para a publicação de dados baseada em uma potente infraestrutura digital, viabilizando o compartilhamento de dados de forma eficaz para que esses recursos atinjam todo seu potencial de reuso (HENNING *et al.*, 2019a).

Dessa forma, em busca de práticas para fomentar a padronização e melhorar o gerenciamento dos dados de pesquisa é que os princípios FAIR se apresentam como um grupo mínimo de requisitos orientadores, nos quais, afirmam HENNING *et al.* (2019b, p. 398) “devem ser aceitos pela comunidade de produtores e consumidores de dados de pesquisa, com a finalidade de incorporar boas práticas para a publicação e compartilhamento de dados científicos”.

As primeiras manifestações referentes a esses princípios surgiram no início de 2014 quando um grupo de especialistas organizaram uma conferência chamada Jointly Designing a Data FAIRPORT e ficou registrado nesse encontro

a necessidade da criação de uma infraestrutura global que possibilitasse dar suporte às publicações e dados originados de pesquisas, propiciando seu compartilhamento e reutilização.

MONTEIRO e SANT'ANA (2019) reiteram que o FAIR proporciona um conjunto de princípios para melhorar a finalidade dos dados e marca um refinamento significativo dos conceitos essenciais para dar-lhes maior valor e ampliar sua capacidade para reutilização, seja por máquinas ou seres humanos, ajudando na interatividade entre aqueles que desejam usar os dados e aqueles que os fornecem.

Os princípios FAIR não propõem que todos os dados sejam livremente disponíveis, principalmente aqueles em que a privacidade deve ser preservada, mas, propõem que os dados estejam disponibilizados para reutilização sob condições e licenças definidas (SCIENTIFIC ELECTRONIC LIBRARY ONLINE, 2016).

## Procedimentos metodológicos

O levantamento bibliográfico no DOAJ, trata-se de um estudo exploratório, utilizando da técnica da Bibliometria.

A FIGURA 1 apresenta as etapas da pesquisa.

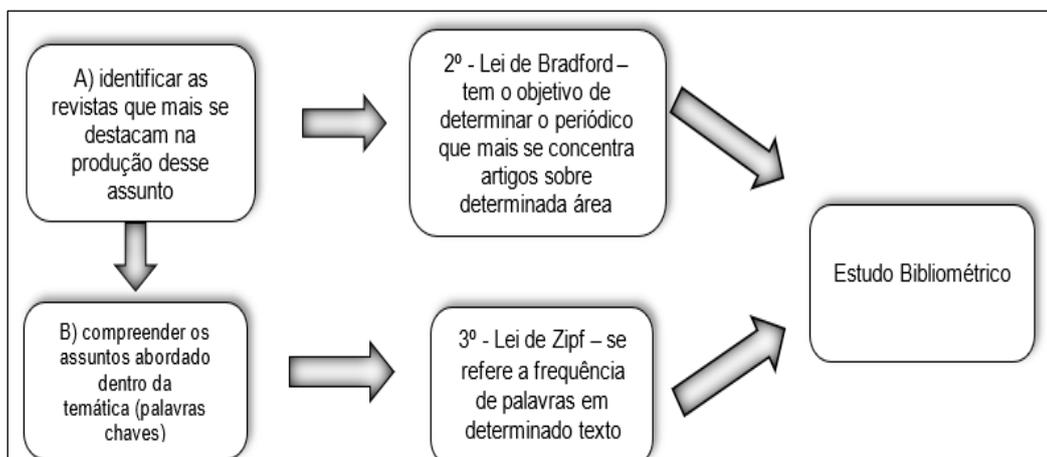


FIGURA 1. Etapas da pesquisa

Fonte: As autoras

O universo de pesquisa consistiu nos periódicos científicos do Directory of Open Access Journals (DOAJ), não houve filtro por área de conhecimento. Foi utilizado o Zotero como ferramenta para gestão dos dados, trata-se de um software livre para gestão de referências bibliográficas de código aberto, gerenciador de citações, arquivador de documentos e ferramenta de colaboração, sendo muito importante no desenvolvimento de trabalhos acadêmicos (SIMONINI, 2019; ZOTERO, 2019).

A coleta de dados aconteceu entre os dias 26 a 30 de agosto de 2021. Os termos de busca no diretório foram “princípios fair” e “fair principles”. Não houve recorte temporal. As palavras-chave escolhidas deu-se pelo objetivo geral deste estudo: investigar a produção científica sobre os Princípios FAIR a partir do Directory of Open Access Journals (DOAJ).

Foram recuperados 103 documentos —97, com o termo em inglês, e 6 em português—, sendo salvos no Zotero para então começar a análise. Posteriormente, foi realizada uma leitura dos trabalhos recuperados a fim de filtrar quais discorrem sobre o tema pesquisado. Constatou-se que, dos 103 documentos recuperados no DOAJ, 40 dissertavam sobre o tema da pesquisa. Os 63 trabalhos descartados durante a análise verificou-se que eles tinham os termos nas palavras-chave, resumo e até mesmo no corpo do texto, no entanto, o trabalho em si não discorria sobre os FAIR, apenas citava-os superficialmente, a discussão não tratava sobre os princípios, apenas mencionava-os. Dessa forma, o corpus de análise foram 40 trabalhos. No próximo momento, procedeu-se a análise quantitativa, para fins de cumprir com os objetivos específicos a) e b).

## Resultados

Conforme proposto pelos objetivos desta pesquisa, inicialmente identificou-se as revistas que mais publicaram a respeito do assunto. O Gráfico 1 ilustra os resultados recuperados.

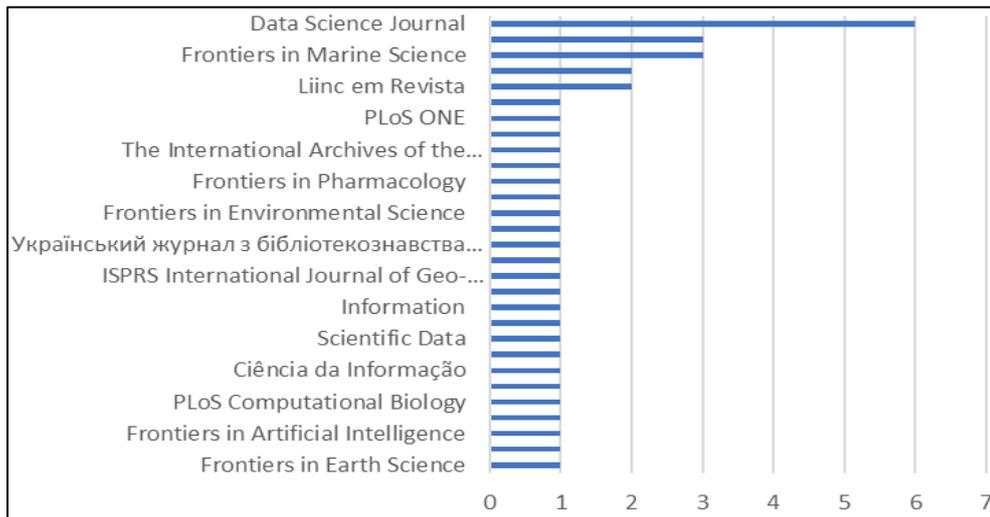


GRÁFICO 1. Revistas que mais publicaram sobre o tema

Fonte: Dados da pesquisa

O GRÁFICO 1 ilustra a revista que mais se destacou em publicações sobre o assunto é a Data Science Journal com 6 trabalhos publicados. Ela é um periódico eletrônico localizado na Inglaterra, revisado por pares, de acesso aberto, que publica artigos sobre a gestão, disseminação, uso e reutilização de dados de pesquisa e bancos de dados em todos os domínios de pesquisa, incluindo ciência, tecnologia, humanidades e artes (DATA SCIENCE JOURNAL, 2021).

Posteriormente, vem o periódico Frontiers in Marine Science, com sede na Suíça, com 3 publicações, que dá visibilidade a pesquisas de todos os aspectos do meio ambiente, biologia, funcionamento do ecossistema e interações humanas com os oceanos (*Frontiers in Marine Science*, 2021.) Ainda, temos a International Journal of Digital Curation, também da Inglaterra, com 3 publicações na qual seus artigos abordam aspectos políticos, estratégicos, operacionais, experimentais, de infraestrutura ou baseados em ferramentas da gestão e curadoria de dados digitais e outros objetos de valor para a pesquisa, patrimônio cultural ou sociedade (*International Journal of Digital Curation*, 2021).

As revistas Liinc em Revista e PeerJ Computer Science aparecem com 2 publicações cada. O primeiro periódico trata-se de uma publicação brasileira de acesso aberto e avaliado por pares, orientada para a reflexão crítica sobre dinâmicas de produção, circulação e apropriação da informação e do conhecimento, ante as transformações do mundo contemporâneo. A PeerJ Computer Science é um periódico internacional, dos Estados Unidos, que primeira cobre todas as áreas disciplinares em ciência da computação. (LIINC em revista, 2021; PEERJ Computer Science, 2021).

Os dados da pesquisa mostraram que as revistas com maior número de publicações sobre os princípios FAIR, são internacionais, revisadas por pares e de acesso aberto e abrangem diversas áreas do conhecimento. Podemos inferir com isso, que os princípios FAIR são para todos, universalmente, e graças ao movimento do acesso aberto podemos ter acesso aos estudos até agora feitos sobre o assunto e seus resultados que já desenvolveram.

A FIGURA 1 ilustra a localização geográfica das revistas que foram identificadas.

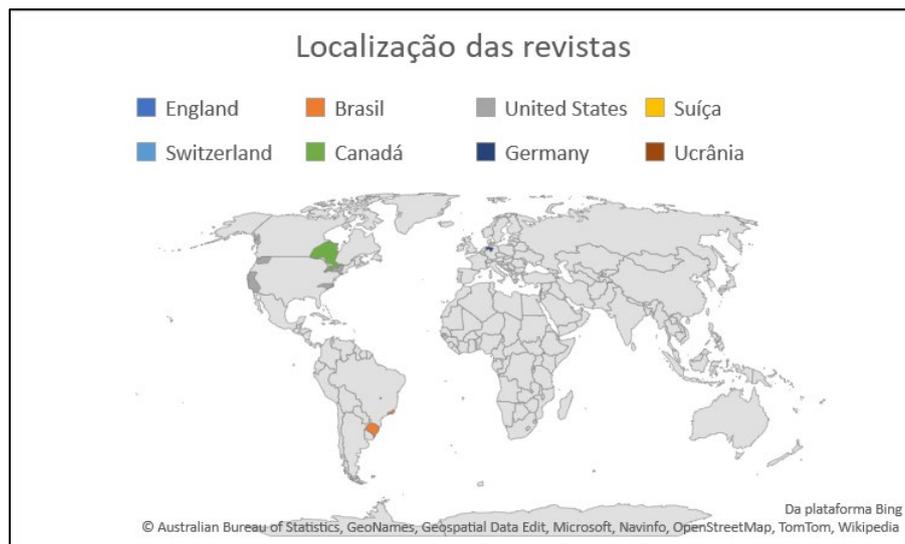


FIGURA 1. Localização geográfica das revistas

Fonte: Dados da pesquisa

Visualiza-se na FIGURA 1 que as regiões que aparecem em cores mais fortes representam as localizações das revistas. Dez revistas pertencem à Suíça, nos distritos de Vaud (distrito de Lausanne) e Basel (distrito de Basiléia). Na Inglaterra oito revistas foram localizadas, pertencem as cidades de: Londres (distrito de London boroughs), Bath (distrito de Somerset), York (distrito de condado de Yorkshire) e Chester (distrito de condado de Cheshire). Seis revistas foram identificadas nos Estados Unidos da América, nas cidades de Chapel Hill (distrito de Carolina do Norte), Nova York (distrito de Nova York), San Francisco (Califórnia) e Washington D. C. (distrito de Washington). Três revistas estão no território brasileiro, a saber, Brasília (Distrito Federal), Rio de Janeiro (Rio de Janeiro) e Porto Alegre (Rio Grande do Sul). Uma revista foi localizada no Canadá, em Toronto (distrito de Ontário). Uma revista está na Alemanha, em Göttingen (Baixa Saxônia). Por fim, uma revista foi localizada na Ucrânia, cidade de Kiev.

O segundo objetivo foi identificar os assuntos mais abordados, ao analisar os artigos recuperados, a FIGURA 2 apresenta uma nuvem de palavras com os temas mais abordados.



FIGURA 2. Nuvem de palavras conforme a ocorrência de termos

Fonte: Dados da pesquisa

A FIGURA 2 mostra os temas que se destacaram por meio das palavras-chave. A partir das sentenças empregadas na busca de pesquisa acreditou-se que a temática mais recorrente seria princípios FAIR, uma vez que entre os objetivos do estudo busca-se conhecer os assuntos abordados. No entanto, conforme a figura 3, vê-se que as palavras em destaque são: Princípios FAIR; Dados FAIR; Repositório de Dados; Ciência Aberta; FAIR; Gestão de dados; interoperabilidade; Infraestrutura de pesquisa e Gerenciamento de dados de pesquisa.

Pode-se compreender dessa forma que os trabalhos dispostos para análise concentram seus assuntos e temas não só sobre os princípios FAIR mas também em tópicos que abrangem um conceito “guarda-chuva” para o estabelecimento de metas e avaliação do sucesso para o gerenciamento de dados eficaz para a comunidade de pesquisa global que empregam os princípios FAIR de modo geral (EASTERDAY *et al.*, 2018).

## Considerações finais

As discussões sobre os princípios FAIR vieram com o intuito de ajudar a gerenciar grandes coleções de dados heterogêneos e reconhecerem que os princípios da ciência de dados podem ajudar pesquisadores em seu trabalho. Essa estrutura é um dos vários esforços recentes que visa estabelecer as melhores práticas para o gerenciamento de dados, eficaz pela comunidade de pesquisa global.

Com as produções científicas destaca-se cada vez mais os grandes volumes de dados que surgem na resolução de problemas persistentes e complexos, levando uma análise recente a declarar os dados como o "recurso mais valioso do mundo".

Diante disso, este trabalho teve como objetivo investigar a produção científica sobre os Princípios FAIR a partir do Directory of Open Access Journals (DOAJ). Para cumprir tal proposta foram alinhados objetivos específicos para dar seguimento à pesquisa. O objetivo a) propôs identificar as

revistas que mais publicaram sobre o tema, e conforme apresentada no estudo notou-se que os periódicos *Data Science Journal*, *Frontiers in Marine Science*, *International Journal of Digital Curation*, *Liinc em Revista* e *PeerJ Computer Science* se destacaram.

Em relação ao objetivo específico b), conhecer os assuntos abordados dentro da temática (palavras chaves), verificou-se que os assuntos destacados foram Princípios FAIR; Dados FAIR; Repositório de Dados; Ciência Aberta; FAIR; Gestão de dados; interoperabilidade; Infraestrutura de pesquisa e Gerenciamento de dados de pesquisa.

Espera-se com essa pesquisa possibilitar mais discussões acerca dos princípios FAIR, uma vez que o gerenciamento e o compartilhamento eficazes de dados são essenciais para a descoberta de conhecimento.

## Bibliografía

- ALBAGLI, SARITA. (2017). Ciência aberta como instrumento de democratização do saber. *Trab. Educ. Saúde*, 15(3): 659-664. <https://www.scielo.br/j/tes/a/NjMd4mXhD43CKqXPcZKrmjj/?format=pdf&lang=pt>
- AGÊNCIA USP DE GESTÃO DE INFORMAÇÃO ACADÊMICA - AGUIA (2022). Entenda o que é Acesso Aberto. [Portal eletrônico da USP]. <https://www.aguia.usp.br/apoio-pesquisador/aceso-aberto-usp/entenda-o-que-e-aceso-aberto/>
- ARAÚJO, CARLOS ALBERTO. (2006). Bibliometria: evolução histórica e questões atuais. *Questão*, 12 (1): 11-32. <https://seer.ufrgs.br/EmQuestao/article/view/16/5>
- BIAZON, TÁSSIA; MARIN, THAIS (2016, julho 22). Ciência aberta: uma nova forma de fazer ciência. (Site eletrônico). <https://www.dicyt.com/noticia/ciencia-aberta-uma-nova-forma-de-fazer-ciencia>
- BRANDÃO, TIAGO; MOREIRA, AMILTON; TANQUEIRO, SARA RAMALHO. (2021). As políticas de acesso aberto: história, promessas e tensões. *Ler História*, 78: 253-276. <https://journals.openedition.org/lerhistoria/8560>

- CAFÉ, LIGIA MARIA ARRUDA; BRASCHER, MARISA. Organização da informação e bibliometria. (2008) *Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf.*, n. esp. <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2008v13nesp1p54>
- COSTA, MICHELLI PEREIRA DA, LEITE, FERNANDO CÉSAR LIMA. (2016). Acesso Aberto no mundo e na América Latina: uma revisão a partir da BOAI. *Transinformação*, 28(1): 33-46. [https://repositorio.unb.br/bitstream/10482/19433/2/AcessoAbertoMundo\\_%202016.pdf](https://repositorio.unb.br/bitstream/10482/19433/2/AcessoAbertoMundo_%202016.pdf)
- COSTA, SELY MARIA DE SOUZA. (2006). Filosofia aberta, modelos de negócios e agências de fomento: elementos essenciais a uma discussão sobre o acesso aberto à informação científica. *Ciência da Informação*, 35(2): 39-50. <https://www.scielo.br/j/ci/a/XsgXnnC7xWHNR7gXrP9Hw3M/?format=pdf&lang=pt>
- CONHEÇA O ZOTERO. <https://www.zotero.org/>
- Data Science Journal*. About this Journal. (2021). <https://datascience.codata.org/>
- EASTERDAY, KELLY *et al.* (2018). From the Field to the Cloud: A Review of Three Approaches to Sharing Historical Data From Field Stations Using Principles From Data Science. *Front. Environ. Sci.*, 6:1-13. <https://www.frontiersin.org/articles/10.3389/fenvs.2018.00088/full>
- Frontiers in Marine Science*. Scope & Mission (2021). <https://www.frontiersin.org/journals/marine-science>
- GUMIEIRO, KATIUCIA ARAUJO; COSTA, SELY MARIA DE SOUZA. (2012). O uso de modelos de negócios por editoras de periódicos científicos eletrônicos de acesso aberto. *Perspectivas em Ciência da Informação*, 17(4): 100-122 <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/1489/1069>
- HENNING, PATRÍCIA CORRÊA; RIBEIRO, CLÁUDIO JOSÉ SILVA; SALES, LUANA FARIAS; MOREIRA; JOÃO LUIZ REBELO; SANTOS, LUIZ OLAVO BONINO. (2019a). Desmistificando os princípios FAIR: conceitos, métricas, tecnologias e aplicações inseridas no ecossistema dos dados FAIR. *Pesq. Bras. em Ci. da Inf. e Bib.*, 14(3): 175-192. <https://periodicos.ufpb.br/index.php/pscib/article/view/46969>
- HENNING, PATRÍCIA CORRÊA; RIBEIRO, CLÁUDIO JOSÉ SILVA; SANTOS, LUIZ OLAVO BONINO; SANTOS, PAULA XAVIER DOS. (2019b.) GO FAIR e os princípios FAIR: o que representam para a expansão dos dados de pesquisa no âmbito da Ciência Aberta. *Questão*, 25(2), 389-412. <https://seer.ufrgs.br/EmQuestao/article/view/84753>
- LIINC em Revista*. Sobre a Revista. (2021). <http://revista.ibict.br/liinc/about>

Manual de formação em ciência aberta. (2018). Conceito e Princípios da Ciência Aberta.

<https://foster.gitbook.io/manual-de-formacao-em-ciencia-aberta/>

MONTEIRO, ELIZABETE CRISTINA DE SOUZA DE AGUIAR, SANT'ANA, RICARDO CÉSAR GONÇALVES.

(2019). Repositórios de Dados Científicos na Infraestrutura de Pesquisa: adoção dos princípios FAIR. *Ci. Inf.*, 48 (Supl.): 347-353.

<http://revista.ibict.br/ciinf/article/view/4878>

PEERJ Computer Science (2021). Who are we? <https://peerj.com/about/team/>

QUEIROZ, CLAUDETE FERNANDES DE, ARAUJO, LUCIANA DANIELLI DE, NASCIMENTO, ANDRÉA

GONÇALVES DO. (2019). Acesso Aberto e Preservação Digital: atuação dos Repositórios Institucionais neste contexto. Semana Internacional de Acesso Aberto, Rio de Janeiro, RJ, Brasil, 1.

[https://www.arca.fiocruz.br/bitstream/icict/36773/2/Apresentacao\\_Semana\\_Internacional\\_AA\\_Preservacao\\_Digital\\_Claudete\\_Queiroz.pdf](https://www.arca.fiocruz.br/bitstream/icict/36773/2/Apresentacao_Semana_Internacional_AA_Preservacao_Digital_Claudete_Queiroz.pdf)

RODRIGUES, CHARLES; VIEIRA, ANGEL FREDDY GODOY. (2016). Estudos bibliométricos sobre a

produção científica da temática Tecnologias de Informação e Comunicação em bibliotecas. *InCID: R. Ci. Inf. e Doc.*, 7(1): 167-180.

<https://www.revistas.usp.br/incid/article/view/98761>

SANTOS, PAULA XAVIER DOS; ALMEIDA, BETHÂNIA DE ARAÚJO; HENNING, PATRÍCIA (orgs.). (2017)

Livro Verde - Ciência aberta e dados abertos: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional. Rio de Janeiro: Fiocruz.

<https://www.arca.fiocruz.br/bitstream/icict/24117/2/Livro-Verde-07-06-2018.pdf>

SCIENTIFIC ELECTRONIC LIBRARY ONLINE. (2016). FAIR guiding principles published in journal of the Nature Publishing Group family. *SciELO in Perspective*, [s.n].

<https://blog.scielo.org/blog/2016/03/16/principios-orientadores-fair-publicados-em-periodico-do-nature-publishing-group/#.X6oQh2hKjDf>

SIMONINI, LEONARDO. O que é o Zotero? (2019) (Blog).

[http://planetazotero.blogspot.com/p/blog-page\\_29.html](http://planetazotero.blogspot.com/p/blog-page_29.html)

# Políticas nacionais de apoio à ciência aberta: análise contrastiva das iniciativas desenvolvidas nos Estados Unidos, Portugal e Brasil

## **IZABEL ANTONINA DE ARAÚJO**

Universidade Federal de Minas Gerais (UFMG) (Brasil)

[antoninaizabel@gmail.com](mailto:antoninaizabel@gmail.com)

## **MARIA APARECIDA MOURA**

Universidade Federal de Minas Gerais (UFMG) (Brasil)

[cidamoura@gmail.com](mailto:cidamoura@gmail.com)

### **RESUMO**

Este trabalho apresenta elementos estruturantes de políticas de ciência aberta e indica como as iniciativas de abertura da ciência têm sido constituídas nos Estados Unidos, Portugal e Brasil. Trata-se de uma pesquisa qualitativa interpretativa, de natureza aplicada que identifica e analisa as condicionantes de uma política de ciência aberta e os elementos que a sustentam. Utiliza-se como método de investigação o estudo de caso cross case para identificar e analisar as variáveis que permeiam o campo político dos países em estudo. Identifica que as políticas se encontram em grau de evolução diferenciado entre os países. Nos Estados Unidos, observa-se forte envolvimento de atores políticos e sociais nas ações de ciência aberta com várias iniciativas que envolvem o acesso às publicações e aos dados de pesquisa, educação aberta e ciência cidadã. Portugal apresenta uma estrutura consolidada para acesso aberto às publicações, conta com o apoio de atores sociais e políticos, alinhamento com as políticas da União

Europeia, além de trabalhar para aprovação de uma política nacional de ciência aberta no país. O Brasil apresenta avanços importantes nas questões de acesso às publicações, mas ainda não dispõe de um marco legal, uma política governamental orientadora para as agências de fomento à pesquisa, universidades e instituições de pesquisa. Conclui-se que as políticas de ciência aberta com foco no acesso justo, democrático e igualitário pautam-se pelo contexto sociocultural do país e sua efetividade depende da governança de recursos informacionais em rede com ações coordenadas entre os diferentes atores sociais, econômicos e políticos.

#### **PALAVRAS-CHAVE**

Políticas de acesso aberto; Ciência Aberta; governança em rede; acesso à informação científica.

## **1. Introdução**

O acesso à informação é um direito social fundamental e desempenha papel estratégico no desenvolvimento de um país. A extensão desse acesso aos resultados de pesquisa, sem barreiras financeiras, legais e técnicas, pode promover benefícios econômicos, sociais e educacionais que têm sido amplamente reconhecidos pela comunidade científica, por decisores políticos e pela sociedade em geral. Esse reconhecimento é evidenciado na Recomendação da Unesco aprovada por 193 países que concordaram em seguir princípios orientadores e adotar a cultura e a prática de ciência aberta. São iniciativas apoiadas em um conjunto de valores compartilhados, como qualidade e integridade, benefício coletivo, equidade e justiça, diversidade e inclusividade (UNESCO, 2021).

Nesse contexto de abertura do conhecimento científico, a colaboração e o compartilhamento da pesquisa são enfatizados como meios para superar os desafios sociais globais, a exemplo da pandemia da Covid-19, que evidencia, para a ciência, para os governantes e para a sociedade, o papel crucial do acesso aos dados de pesquisa. Em razão disso, intensifica-se ainda mais o movimento pela ciência aberta que, além do acesso às publicações, pleiteia a abertura de metodologias, códigos, softwares, dados de pesquisas, revisão por pares aberta e transparente, ciência cidadã, entre outras formas de abertura (ALBAGLI, 2015). Este tema é amplamente defendido em termos de reprodutibilidade, disseminação e transferência de novos conhecimentos para a sociedade tendo em vista a ampliação das possibilidades de inovação, não apenas em pesquisas científicas, mas também na forma de comunicar a ciência (UNESCO, 2021).

A ciência aberta é definida pela Unesco como um construto inclusivo que combina vários movimentos e práticas de disponibilizar abertamente o conhecimento científico. As ações nesta direção são orientadas a fomentar o compartilhamento para além da comunidade científica, visando promover a inclusão e o intercâmbio do conhecimento acadêmico “[...] originário de grupos tradicionalmente sub-representados ou excluídos (como mulheres, minorias, indígenas, acadêmicos de países menos favorecidos e línguas com poucos recursos) e contribuir para reduzir as desigualdades no acesso ao desenvolvimento científico, às infraestruturas e às capacidades entre diferentes países e regiões” (UNESCO, 2021, p. 5).

Esse processo de abertura da ciência tem sido desenvolvido por alguns países com a implementação de políticas de informação específicas que orientam e regulamentam o acesso aos resultados de pesquisas produzidas com recursos públicos. As iniciativas nessa direção têm sido desenvolvidas de forma diferenciada entre os países, tendo em vista que dependem de um conjunto de fatores, tais como: ambiente político favorável, apoio governamental, tecnológico, econômico, social e cultural, além de outros que

podem estar relacionados à cultura política e ao grau de consolidação da democracia, que definem o relacionamento dos cidadãos com o seu governo.

No ambiente de responsabilização política, estão incluídos ainda atores como agências financiadoras de pesquisa, universidades e instituições de pesquisa, sociedade, empresas públicas e privadas. Essa corresponsabilização é importante, pois a ciência aberta requer políticas públicas estruturantes que perpassam o campo científico, como aquelas voltadas para a educação, a cultura e a inclusão social e digital.

Nesse cenário de abertura, países como Estados Unidos e Portugal vêm intensificando várias iniciativas e políticas de acesso aberto à produção científica consubstanciadas em um modelo de governança que contribui para ampliar as políticas de ciência aberta. A experiência brasileira é marcada por várias iniciativas, tendo conseguido avanços consideráveis. Entretanto, as ações não se consolidam efetivamente, em função da ausência de políticas governamentais para fortalecer e articular as políticas institucionais existentes, apoio e incentivos das agências financiadoras de pesquisa para concretizar as ações e traçar um caminho mais célere e sólido para a ciência aberta no país.

Considerando a recomendação da Unesco sobre a necessidade de fomentar um ambiente político favorável para a ciência aberta, este artigo apresenta um recorte de uma pesquisa mais ampla realizada em nível de doutorado. Resulta de um estudo contrastivo das iniciativas e políticas de acesso aberto e ciência aberta nos Estados Unidos, Portugal e Brasil e os seus elementos estruturantes. Na pesquisa de origem, foram observadas as dimensões políticas, sociais, técnicas, tecnológicas e econômicas. Especificamente neste artigo, apresenta-se a centralidade da dimensão política por se considerar esta variável determinante para as demais dimensões, a saber: sociais, técnica, tecnológica e econômica.

Utilizou-se como procedimento metodológico a análise de caso comparativo com base em estudo documental e entrevistas com lideranças de instituições pioneiras em ações de acesso aberto, como Universidade do Minho, SPARC e IBICT. A sistematização do amplo entendimento da implementação das

iniciativas e políticas de acesso aberto dos três países contribuiu na compreensão sobre esse novo cenário de mudança no processo de abertura do conhecimento científico, além de fomentar o debate sobre a importância do engajamento dos atores políticos, econômicos e sociais para ampliar as ações de ciência aberta.

## **2. Políticas de ciência aberta: desafios e perspectivas**

A informação científica é um recurso fundamental para a inovação tecnológica, a expansão do acesso e do compartilhamento do conhecimento, tendo como perspectiva a ampliação de igualdade de oportunidades no que se refere ao uso da informação. “É a base do processo científico, motor do desenvolvimento social, econômico e humano” (SILVA, 2017, p. 8). O acesso ao conhecimento nivela a comunicação científica e o acesso a ela, e seu impacto ocorre além do meio acadêmico, em âmbitos empresarial, de saúde ou educação (SWAN, 2012).

A abertura do conhecimento envolve a relação com a formulação de políticas públicas, a ampliação do interesse e a conscientização do cidadão sobre o trabalho da ciência em prol do progresso econômico, a justiça social, ambiental e cognitiva. Nesta direção, uma ciência aberta de forma compartilhada passa a ser defendida por autores de várias escolas de pensamentos, dentre elas, a escola democrática e a escola pública que trazem uma proposta de repensar o modo de produção da ciência e sua relação com a sociedade (FECHER e FRIESEKE, 2013). Os defensores destas proposições estão preocupados em expandir a base social e a interlocução da ciência com outros tipos de saberes no enfrentamento aos grandes desafios globais.

Em nível governamental, essa abertura é muito difundida pela União Europeia, pelos governos dos Estados Unidos, e de Portugal. Ambos os países vêm desenvolvendo suas políticas buscando proporcionar maior abertura no processo científico com a implementação de propostas que se orientam à ideia de justiça social, conhecimento para todos e com todos. Em defesa dos países

em desenvolvimento, REILLY (2016) advoga sobre a necessidade se compreender a centralidade exercida pela falta de uma política governamental, maior institucionalização das políticas existentes, apoio e incentivos das agências financiadoras de pesquisa para “ajudar pesquisadores e produtores de conhecimento a entender melhor o problema de gerar espaços nos quais há respeito pelos direitos de comunicação, democratização conhecimento e comunicação aberta” (REILLY, 2016, p. 56).

Democratizar o acesso de forma justa e promover a cidadania exige repensar nas novas relações de poder e saber na sociedade contemporânea, e, principalmente nas assimetrias sociais, econômicas e políticas que se observam em âmbito nacional e internacional. Sabe-se que “uma ciência situada em que o conhecimento está estabelecido dentro de relações históricas, políticas e socioculturais particulares com desigualdades e hierarquias de produção de conhecimento e seus conflitos inerentes” (REILLY, 2016, p. 56). Corrigir as injustiças cognitivas requer considerar que todos os indivíduos e comunidades, independentemente de sua cultura, gênero, status socioeconômico ou idioma, devem ser capazes de explorar plenamente suas capacidades de usar, compartilhar e criar conhecimento (JORANSON, 2008). As políticas de informação, que orientam o acesso aberto e ciência aberta, envolvem instâncias de ação e decisão diferenciadas, internas e externas, são geralmente estabelecidas por instituições que realizam pesquisas, financiadores de pesquisas e governos. Em geral, se estabelecem em formato de leis, regulamentos ou orientações em interfaces com diferentes campos do conhecimento podendo ser configuradas como políticas públicas de Estado e ou uma política de governo. No caso da ciência aberta –que é um campo amplo e complexo que interconecta outros campos, como o educacional, o econômico, o tecnológico e o social, demanda interesse e ações de diferentes atores–, requer-se uma política de Estado para garantir a permanência das ações e o apoio em estratégias que garantam o acesso universal ao conhecimento científico. É o “Estado que exerce sua autoridade para dirigir a sociedade, pois detém das condições financeiras e administrativas adequadas

para implementar políticas públicas a serem formuladas e avaliadas socialmente” (RODRIGUES, 2018, p. 22). Além disso, é o agente que mais investe e catalisa recursos em inovação, dinamiza a capacidade e a disposição do empresariado, criando os espaços e as perspectivas de investimento e mercado com incentivos a participação de atores econômicos e científicos em diversas áreas (MAZZUCATO, 2018).

A efetividade das políticas de ciência aberta depende do envolvimento de ações coordenadas entre vários agentes (órgãos públicos, privados, agente políticos, cidadãos, organizações públicas e demais entidades). Depende também de informações precisas sobre o público-alvo e os possíveis impactos provocados nas diferentes camadas sociais. Outros aspectos a serem considerados em uma política de ciência aberta é que as instituições formais e processos de governos geopoliticamente reconhecidos não operam isoladamente (BRAMAN, 2011). Isso implica pensar no envolvimento dos setores públicos e privados que interferem diretamente na aplicação desta lei como: as regras ditadas pelo mercado editorial, a lei de direitos autorais, as condições de acesso e uso da informação pela sociedade (educação, literacia científica, inclusão digital, entre outras).

Os desafios relacionados à formulação de políticas de dados digitais são de diversas ordens e envolvem aspectos técnicos, legais e éticos de como assegurar a heterogeneidade e interoperabilidade dos dados, a acessibilidade, pesquisabilidade, preservação, curadoria, qualidade, avaliação e, também, a segurança dos dados (RECODE, 2014). Outros desafios que são observados nesta direção é que os governos necessitam desenvolver políticas de abertura dos dados científicos que contemplem também as políticas de abertura e o acesso aos dados públicos, bem como apoiar a criação de software e a formação de recursos humanos, elementos cruciais ao sucesso de ambas as políticas (THE ROYAL SOCIETY, 2012).

Estas políticas devem prever a capacitação tanto dos profissionais de informação quanto dos cientistas para o uso dessas ferramentas. A familiaridade com ferramentas e princípios de gestão de dados deve ser,

segundo a Royal Society (2012), parte integrante da formação dos cientistas. Além desse segmento, é necessário direcionar o olhar para os cidadãos, pois envolve engajamento do público e da sociedade civil nas atividades de pesquisa científica com o objetivo de contribuir com seu esforço intelectual, seu conhecimento local ou com suas ferramentas e recursos para o desenvolvimento científico (ALBAGLI, 2015).

Diante das questões apresentadas e sob o entendimento de que no contexto da ciência aberta, é a variável política que pode subsidiar as demais ações sociais, técnicas, tecnológicas e econômicas, apresenta-se nas próximas seções uma síntese das ações políticas associadas à ciência aberta nos Estados Unidos, Portugal e Brasil.

### **3. Metodologia**

Trata-se de um estudo exploratório e descritivo, de natureza qualitativa, que utiliza como técnicas de investigação a pesquisa documental e entrevista semiestruturada. Por meio do estudo de caso (cross-case), foram identificadas e analisadas as iniciativas políticas de acesso aberto e ciência aberta nos Estados Unidos, Portugal e Brasil.

Na pesquisa documental, os principais documentos considerados foram o Plano de Ação Nacional em Governo Aberto – Open Government Partnership (OGP), leis, decretos, declarações e planos emitidos pelos governos destes países na formulação e execução de políticas de acesso aberto à informação.

A entrevista permitiu coletar informações adicionais com as lideranças da Universidade do Minho (Portugal), SPARC (Estados Unidos) e IBICT (Brasil) sobre a elaboração das iniciativas de acesso à informação científica. Possibilitou investigar como esses atores sociais do movimento de acesso aberto veem os processos e fatores que estruturam uma política nacional de ciência aberta. A escolha desses líderes está relacionada a vasta experiência e engajamento com as políticas de acesso aberto e ciência aberta em âmbito nacional e internacional.

PAÍS	INSTITUIÇÃO/CARGO	ATUAÇÃO/AÇÕES
Estados Unidos	Scholarly Publishing and Academic Resources Coalition (Sparc)  Diretora executiva	Atua na organização de bibliotecas, faculdades e universidades nos Estados Unidos e no Canadá. Participação ativa em comitês e projetos de várias agências federais dos Estados Unidos e parte do Conselho Consultivo de Dados de Comércio. Promove ações coletivas de coalizões com vários grupos de trabalho, entre os quais, a ATA e o Open Access Working Group, universidades, bibliotecas, advocacia e grupos de consumidores que atuam como vozes líderes nas políticas de acesso aberto, incluindo a política de acesso público (National Institutes of Health – NIH) e uma diretiva da Casa Branca
Portugal	Universidade do Minho  Diretor do Centro de documentação da Universidade do Minho e Coordenador da equipe técnica da Universidade	Atuou na Criação do RepositórioUM e no projeto de Repositório Científico de Acesso Aberto de Portugal (RCAAP). Promove o acesso aberto e criação de repositórios institucionais em Portugal e nos países lusófonos. Atua no nível europeu como membro do Grupo de Trabalho sobre Acesso Aberto da Associação Europeia das Universidades – European University Association (EUA) –, representando o Conselho de Reitores de Portugal. Coordena a participação da Universidade do Minho em vários projetos, como o Digital Repository Infrastructure (DRIVER), Open Access Infrastructure for Research in Europe (OpenAIRE), NECOBELAC
Portugal	Chefe de Divisão dos Serviços de Documentação da Universidade do Minho e Gerente de Projetos de Ciência Aberta	Participa de vários projetos relacionados ao acesso aberto/ciência aberta e repositórios como: NECOBELAC, OpenAIRE, MedOANET, OpenAIREplus e PASTEUR40A, FIT4RRI e FOSTERplus. Participa do grupo de trabalho COAR sobre interoperabilidade de repositórios e vocabulários controlados. Integra o grupo de trabalho nacional sobre Política Portuguesa de Ciência Aberta

PAÍS	INSTITUIÇÃO/CARGO	ATUAÇÃO/AÇÕES
Brasil	Coordenador do Grupo de Trabalho de Bibliotecas Acadêmicas (Ibict) e do Laboratório de Metodologias de Tratamento e Disseminação da Informação COLAB	Representa o órgão no Brasil e no exterior em várias ações de acesso aberto e ciência aberta. Desenvolve trabalhos relacionados a repositórios digitais científicos que contribuem para o Ibict ser referência em transferência de tecnologias da informação, como a coleta automática de registro e disseminação de teses e dissertações, a editoração de revistas eletrônicas e os repositórios de documentos digitais de diversas naturezas

QUADRO 1. Lista dos entrevistados com respectivos cargos e atuações. Fonte: Elaborado pela autora

Para análise dos dados foi feito o cruzamento das informações delineando as semelhanças e especificidades no contexto dos três países. Conforme YIN (2015, p. 19), “os estudos de caso podem cobrir casos múltiplos e, então, tirar um conjunto único de conclusões de ‘cross-case’”. Para isso, foram criadas categorias analíticas de acordo com as variáveis políticas para subsidiar a compreensão de ações de governo que contribuem para o processo de democratização da informação e ampliação do acesso aos resultados de pesquisas financiadas com recursos públicos. Tais categorias são:

- a)** o comprometimento do governo em ações de ciência aberta;
- b)** a existência de aparato legislativo de acesso à informação científica e aos dados de pesquisa;
- c)** o engajamento político das organizações ligadas à informação e cultura nos referidos países;
- d)** o aporte político das agências de financiamento à pesquisa.

Conforme afirma BARDIN (2011), a categorização é uma operação de elementos constitutivos de um conjunto por diferenciação e, em seguida, por reagrupamento com os critérios previamente definidos. Com base nas categorias definidas, analisaram-se as entrevistas, classificando-se os elementos mais significativos contidos em cada resposta, de forma a ratificar e

complementar os dados obtidos na etapa anterior. Esse procedimento foi sustentado em BARDIN (2011, p. 48), que tangencia a constituição, a forma de operação e os objetivos da análise de conteúdo. Os resultados foram indicados nos Quadros 1, 2, 3 e 4, que fazem a comparação das ações políticas realizadas nos três países estudados.

## **4. Análise e discussão de resultados**

Apresentamos a seguir uma análise das principais ações de acesso aberto desenvolvidas pelos países em estudo e sinalizamos para a necessidade de direcionar o olhar para o contexto sociocultural, que em grande medida influencia o desenvolvimento de ações de ciência aberta.

### **4.1 Orientação política de acesso aberto nos Estados Unidos**

O acesso público à informação é uma prática consolidada há mais de 50 anos nos EUA. A Lei de Liberdade da Informação (FOIA) foi instituída em 1966, o instrumento estabelece que qualquer cidadão tem o direito de fazer solicitações por escrito de registros de agências federais e que as agências precisam fornecer as informações solicitadas, a menos que se enquadre na classe de informações sigilosa (Estados Unidos, 1966).

O autogoverno, que tem como base a ideia de que os cidadãos têm o direito e o dever de participar, contribui enormemente na formação de grupos organizados para defender diversos interesses sociais. Esta coletividade está presente em representações de diversos segmentos públicos e privados. O que caracteriza estes grupos é a alta capacidade de organização para aumentar as chances de serem ouvidos pelos formuladores de políticas (Patterson, 2018). Ações podem ser evidenciadas pelo comprometimento da sociedade americana, principalmente da sociedade científica, de organizações governamentais e não governamentais (universidades, bibliotecas, associações de alunos e professores) em prol da defesa do acesso aberto, cujas ações definiram os rumos das políticas públicas governamentais no país e em âmbito mundial. Dentre as organizações com maior destaque na força de coalizão pelo

acesso aberto, cita-se a Association of Research Library (ARL), que congrega 124 bibliotecas de instituições de pesquisa americanas, A Alliance for Taxpayer Access ATA, formada por 97 instituições representativas de associações científicas, de pesquisadores e cidadãos americanos e a SPARC, com 210 membros, principalmente bibliotecas acadêmicas e de pesquisa localizadas nos EUA e no Canadá.

Nesse contexto, observa-se que a mobilização social foi fundamental para que o governo estadunidense se sensibilizasse para a importância do acesso aberto. Esse movimento liderado pelos pesquisadores estadunidenses evoluiu para iniciativas mais amplas, como as adotadas pelo repositório PubMed Central e as políticas mandatárias de suas universidades, que incentivam grande volume de repositórios institucionais. Esse movimento social desenvolve ações para o acesso aberto articulada com os líderes governamentais, com representação tanto no congresso, quanto no senado. Essa conexão contribuiu para a aprovação da primeira política governamental do National Institutes of Health (NIH). A política obriga o depósito das publicações aprovadas revisadas por pares resultantes de pesquisas financiadas no todo ou em parte no repositório temático PMC com período máximo de embargo de 12 meses (Estados Unidos, 2007).

O país tornou-se membro cofundador da OGP em 2011, adotou como estratégia trabalhar tanto os dados públicos de governo quanto os dados de pesquisa no mesmo programa. Para tanto, o governo emitiu uma diretiva instituindo que as agências federais, com investimento anual acima de 100 milhões de dólares, desenvolvessem no prazo de 06 (seis) meses um plano para disponibilizar gratuitamente as publicações revisadas por pares, que recebem financiamento público. Além disso, determinou que as agências distribuíssem um kit que inclui a divulgação das melhores práticas, treinamento, políticas e orientações relacionadas à inovação aberta, incluindo abordagens como prêmios de incentivo à ciência cidadã e o crowdsourcing (Estados Unidos, 2011, 2013).

Concomitante a estas ações, o governo buscou o apoio da sociedade para ajudar as agências de pesquisa a abordar as informações científicas na resolução de problemas sociais. Dentre as iniciativas está a expansão do uso de tecnologia para atingir maior eficiência na administração de informações públicas, aprimoramento da interface de participação pública em sites do governo dentre as quais: a ExpertNet, a “We the People” que capacita o público a criar petições online e a falar diretamente com o governo.

Este movimento em torno da defesa da ciência cidadã contribuiu para que fosse instituída uma seção (402), conhecida como ciência cidadã e Crowdsourcing, na Lei Americana de Inovação e Competitividade (AICA). A lei concede às agências federais ampla autoridade para usar crowdsourcing para avançar em missões das agências e facilitar participação do público no processo de inovação (Estados Unidos, 2017).

Nos Estados Unidos, os dados de pesquisa são tratados como ativo estratégico com ênfase na inteligência artificial. Em 2019, foi sancionada a lei de dados abertos. Esta legislação institucionaliza o compromisso do governo federal com o Open Data e exige que as agências federais publiquem dados do governo em formatos abertos e legíveis por máquina e que façam uso de licenças abertas. Além disso, o governo orienta as agências a apoiarem usos inovadores de dados e a desenvolver melhores práticas para o Open Data. Atualmente, articula ações com financiadores públicos e privados, instituições e universidades para a criação de um plano de gerenciamento de dados que possa garantir infraestrutura para o compartilhamento, sustentabilidade, armazenamento e preservação de dados de pesquisa a longo prazo.

Além das políticas para publicações, dados e ciência cidadã, os Estados Unidos desenvolvem com muito destaque a educação aberta. O objetivo do país é promover a educação e a democratização do acesso ao conhecimento sem barreiras legais, financeiras e técnicas. Em 2018, quase a metade de todos os estados do país adotavam políticas, procedimentos e diretrizes de acessibilidade de materiais didáticos (SPARC, 2020). Em 2020 uma grande força de coalizão de apoiadores do acesso aberto enviou uma carta ao governo

solicitando a atualização da lei de acesso público, com eliminação do período de embargo (SPARC, 2020).

Para além de uma política central bem coordenada, as condições que favorecem os Estados Unidos a atingir esse status de abertura da ciência estão relacionadas com a sua cultura política que é reflexo, em grande parte, da tradição de participação local de dezenas de milhões de participantes se envolvem em assuntos comunitários através de associações de pais, professores, grupos de bairro, de igrejas e hospitais (PATTERSON, 2018). O país é também responsável pelo maior sistema de faculdades do mundo com pioneirismo na ideia de educação pública, no início de 1800. Tem um sistema de ensino superior que inclui mais de três mil instituições de dois a quatro anos de duração do estudo. “Ainda hoje, embora alguns dos jovens não tem uma chance realista de frequentar a faculdade, o sistema de faculdade é relativamente aberto (PATTERSON, 2018, p. 16). Entre adultos com idade entre 20 e 25 anos ou mais, aproximadamente um em cada quatro é formado. Um quarto dos cidadãos adultos do país possui um diploma universitário (PATTERSON, 2018). A FIGURA 1 abaixo descreve a condução das políticas de acesso aberto nos Estados Unidos.

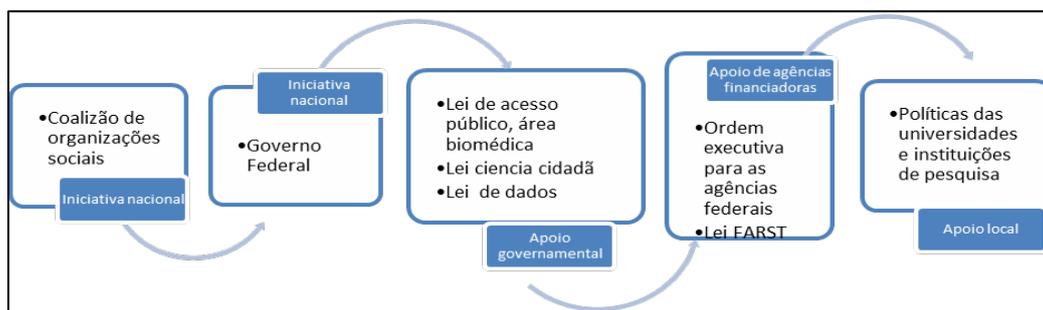


FIGURA 1. Estratégia política de acesso aberto nos EUA

Fonte: Elaborada pelas autoras

Observa-se pela figura que a direção da política de acesso aberto nos Estados Unidos é impulsionada por forte coalizão de organizações (sociedades científicas, e associações representante da sociedade civil, bibliotecas, universidades e pesquisadores). As políticas regulamentadoras em nível federal orientam as ações para as agências do governo, estas por sua vez, estabelecem suas políticas de financiamento para universidades e instituições de pesquisa.

#### **4.2 Orientações políticas de acesso aberto em Portugal**

Portugal tem um protagonismo assinalável na adoção e implementação de políticas destinadas à promoção do acesso aberto e vem se estruturando para a implementação de uma política nacional de ciência aberta. O país apresenta um nível de maturidade bem elevado com relação à disponibilização de publicações científicas. Desde 2003, sob a liderança do centro de documentação da Universidade do Minho, vem implementando políticas e garantindo as infraestruturas necessárias para ampliação do acesso aberto dentre elas: a implementação dos repositórios institucionais, políticas mandatórias de auto arquivamento da produção e criação e gestão do Portal Nacional agregador, RCAAP, que indexa os conteúdos científicos em acesso aberto existentes nos repositórios institucionais das entidades portuguesas de ensino superior, e outras organizações de Pesquisa e desenvolvimento do país (Portugal, 2016).

Um diferencial, que contribui para o sucesso da ampliação de políticas de acesso aberto em Portugal, é a atuação do Conselho de Reitores das Universidades Portuguesas (CRUP). Esta entidade atua na coordenação do ensino universitário e integra, como membros efetivos, o conjunto das universidades públicas e a Universidade Católica Portuguesa. Este Conselho estabeleceu um grupo de trabalho sobre acesso aberto em 2007, que opera na mobilização da comunidade acadêmica nas práticas de acesso aberto. O apoio dos reitores e a posição política diante do governo foi muito importante para o país conseguir dar visibilidade e reconhecimento do portal a nível nacional e internacional.

Outro diferencial na condução das políticas em Portugal é o apoio da principal agência de fomento do país, a Fundação para a Ciência e a Tecnologia (FCT), que definiu a sua política de financiamento em 2014. Soma-se a isso, o alinhamento do país com as políticas da União Europeia. Em 2012, a Comissão Europeia recomendou que os Estados-Membros desenvolvessem políticas nacionais para disponibilizar os resultados de pesquisa financiada com recursos públicos e que os financiadores de ciência e as instituições de pesquisa desenvolvessem as suas próprias políticas, coordenadas em nível nacional e Europeu. As ações políticas em Portugal são trabalhadas junto ao Ministério da Ciência, Tecnologia e Ensino Superior (MCTES). Por meio deste órgão, o governo elegeu a promoção do conhecimento para todos como pilar do seu programa e está empenhado na elaboração e implementação de uma política nacional de ciência aberta.

Em 2016, o governo português publicou uma Resolução nº 21 do Conselho de Ministros no qual são definidos os princípios orientadores e práticas de ciência aberta, estabelece metas a curto e médio prazo para o desenvolvimento de iniciativas: uma delas é a elaboração de uma Carta de Compromisso para a Ciência Aberta em Portugal, envolvendo instituições governamentais, pesquisadores, agências de financiamento de ciência, instituições de ensino superior, unidades de pesquisa, arquivos, bibliotecas, editoras, setor empresarial e organizações de base científica e tecnológica (PORTUGAL, 2016). Outra iniciativa foi a elaboração de um plano estratégico para a implementação da política nacional de ciência aberta. Esta proposta foi formulada com base em estudos realizados por uma Comissão Executiva composta por membros experientes e atuantes nas instituições do sistema científico de informação (bibliotecas e museus), tecnologia e cultura. A representatividade alcançada em tais iniciativas permitiu diagnosticar a necessidade de políticas públicas, bem como identificar os pontos positivos e negativos vinculados à ciência aberta.

Observa-se que a participação da sociedade portuguesa nos assuntos científicos, ainda é baixa, como constatado no relatório do estudo desenvolvido pelo comitê interministerial para ciência aberta realizado em 2016. O estudo atribui esta reduzida participação, ao contexto histórico de baixa literacia científica e pouco interesse dos portugueses pelas discussões políticas envolvendo a temática do desenvolvimento científico e tecnológico. Em virtude de tais constatações, o país vem trabalhando intensamente em questões associadas à responsabilidade social científica através da promoção de capacitação e formação profissional para diferentes segmentos sociais, dentre eles os cidadãos comuns, os cientistas e jornalistas (Portugal, 2016).

Portugal tem uma democracia jovem, viveu longos períodos de restrição de direito. Nesse sentido, para além dos esforços nacionais, o sucesso das políticas no país está associado, em grande medida, ao alinhamento com as políticas da União Europeia que com a potência de um bloco econômico, pauta diretrizes e metas para ajustamento interno dos países membros. Este aporte foi fundamental para implementação de políticas, principalmente as políticas transversais que contribuem para a ciência aberta, como políticas educacionais para elevar o nível de formação da população e promover a literacia científica. Conforme o relatório da EU, o índice de diplomados portugueses no ensino superior subiu de 27,8% em 2012 para 33,5% em 2018 entre alunos de 30 a 34 anos. Este índice está ainda um pouco distante da meta da UE que é de 39,9 (Portugal, 2016). As estratégias adotadas por Portugal na condução das políticas de acesso aberto, está ilustrada na FIGURA 2 abaixo:

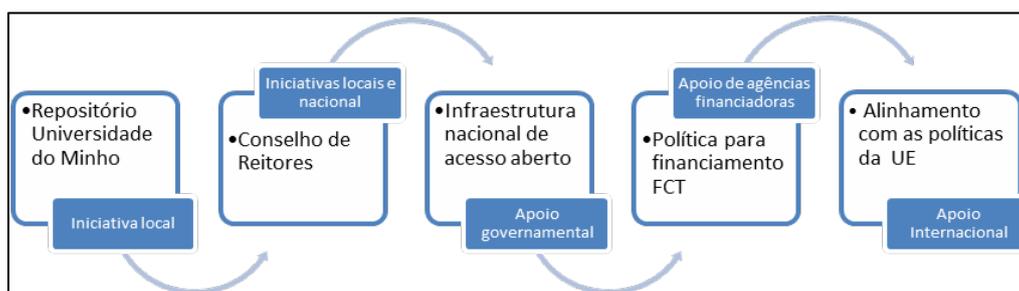


FIGURA 2. Estratégia política de acesso aberto em Portugal

Fonte: Elaborado pelas autoras

Constata-se, pela descrição da FIGURA 2, que Portugal adotou um caminho inverso ao dos EUA, mas estrategicamente com o mesmo objetivo e mesma finalidade; iniciou as ações de acesso aberto no ponto base, ou seja, nas universidades, onde promoveu campanhas de sensibilização da comunidade usuária. Ao longo da trajetória conseguiu envolver paulatinamente diferentes entidades, em ações programadas de efeito cascata. Em princípio, buscou apoio em nível local, no contexto da gestão universitária e envolveu gestores e conselhos. Essa ação política foi fundamental para conseguir aporte de recursos do governo e apoio da FCT no processo de alinhamento internacional com as políticas da UE. Essa estratégia de envolvimento institucional colocou o país em destaque mundial na sensibilização para as políticas de auto arquivamento, tendo em vista que, as pesquisas são realizadas principalmente nas universidades públicas que concentram o maior número de pesquisadores.

#### **4.3 Orientações políticas de acesso aberto no Brasil**

O Brasil já participava de iniciativas de apoio ao acesso aberto à informação científica, mesmo antes da intensificação desse movimento, ocorrido em 2002. Uma de suas primeiras ações foi a concepção do projeto SciELO (Scientific Electronic Library Online) em 1997, iniciativa contribuiu para o país se tornasse referência na publicação de revistas de acesso aberto.

Em 2005, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), formalizou apoio ao acesso aberto, e, diferente dos Estados Unidos e Portugal, investiu em duas estratégias de acesso (via dourada –publicações em revista de acesso aberto e via verde– disponibilização da produção em repositórios institucionais).

Ao longo de duas décadas, vem implementando sistemas importantes para ampliação do acesso aberto no país, dentre eles: Portal brasileiro de publicações científicas em acesso aberto (Oasisbr); Sistema de Publicação Eletrônica de Teses e Dissertações (TEDE) integrado à Biblioteca Digital Brasileira de Teses e dissertações (BDTD); Sistema Eletrônico de Editoração de Revistas (SEER), apoio às instituições para na implementação de repositórios,

dentre outras. O órgão está no processo de desenvolvimento de políticas e infraestrutura para os dados de pesquisas.

O país tornou-se membro cofundador da OGP em 2011, mas ao contrário dos EUA, inicialmente, não priorizou acesso aberto aos resultados das pesquisas científicas. Suas ações buscaram contemplar mais especificamente a disponibilização de informações e dados administrativos geradas em órgãos do poder executivo. O tema ciência aberta apareceu pela primeira vez no marco três do quarto plano Nacional da OGP, em 2018. Desde então, vários órgãos governamentais assumiram um compromisso junto ao comitê interministerial para o desenvolvimento de ações para ampliar a ciência aberta no país. Este cenário é muito promissor com boas perspectivas, pois o grupo de trabalho conta com vários órgãos, dentre os quais a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico.

O governo brasileiro, até o momento, não propôs legislação diretamente relacionada ao acesso aberto e nem formalizou orientações para as agências públicas de fomento à pesquisa. Observa-se que o país carece de apoio destes agentes institucionais. Conforme CARUSO, NICOLE ARCHAMBAULT (2014), a existência de mandatos de depósito obrigatório, sobretudo por parte das agências de fomento, pode conduzir a uma maior consciencialização do acesso aberto e promover maior disponibilidade de artigos científicos em repositórios institucionais. O país possui atualmente 29 principais agências financiadoras de pesquisa, sendo três agências federais e 26 estaduais, mas somente a Fapesp, agência de apoio à pesquisa de São Paulo, instituiu uma política mandatória de acesso aberto.

O Sistema de avaliação da produção científica brasileira ainda não dispõe de uma política de incentivo para publicação em acesso aberto. A política da Capes segue na contramão desse propósito, tendo em vista que essa agência avalia os programas de pós-graduação utilizando principalmente o quantitativo de publicações em revistas ranqueadas por fator de impacto, a maioria delas de acesso restrito. Observa-se que a situação se torna ainda mais complexa

porque a ausência de políticas mandatórias por parte do governo, das agências de fomento e por grande parte das instituições de pesquisa reduz drasticamente as oportunidades para integrar as ações de abertura dos dados de pesquisas afetando, desse modo, as mudanças nos sistemas de avaliação interna de docentes e pesquisadores.

O envolvimento das organizações ligadas à informação, à cultura e ao ensino no país nas ações de ciência aberta ainda é incipiente e, e embora o Brasil tenha conquistado avanços, com a implementação dos repositórios institucionais e bibliotecas digitais, a manutenção e ampliação desses projetos para maior prática de abertura é ainda desafiador devido a ausência de institucionalização das ações, de infraestrutura técnica e ao reduzido quadro de recursos humanos dedicados aos projetos e programas.

O Brasil tem uma democracia jovem com uma tradição cívica pouco encorajadora ao uso de fontes diversificadas de informação em função de heranças históricas e políticas associadas a longos períodos de restrição de direitos que contribuíram no estrutural de diversas ordens. A desigualdade social e o atraso educacional são certamente as piores barreiras ao acesso aberto, pois dificultam as possibilidades de o país atingir os patamares de outras nações em termos de desenvolvimento inclusivo que se comprometa com a ampliação de oportunidades para todos. Estas adversidades impõem grandes desafios ao Brasil no que se refere à formulação e a consolidação de políticas públicas de abertura do conhecimento científico produzido no país.

Esse quadro se agrava, se considerarmos que grande parte da população não percebe que os recursos direcionados para pesquisa, em sua maioria, são provenientes de impostos pagos por ela. Como exemplo, citam-se os dados da PNAD Contínua de 2019 divulgado pelo IBGE, apesar da proporção de pessoas de 25 anos ou mais com ensino médio completo ter crescido no país, passando de 45,0 % em 2016 para 47,4 % em 2018 e 48,8 % em 2019, mais da metade (51,2 % ou 69,5 milhões) dos adultos não concluíram essa etapa educacional. A situação do ensino básico também é muito grave, pois 50 milhões de pessoas de 14 a 29 anos do país, 20,2 % (ou 10,1 milhões) não

completaram alguma das etapas da educação básica, seja por terem abandonado ou não terem frequentado a escola antes do término desta etapa, seja por nunca a ter frequentado. “Nesta situação, portanto, havia 10,1 milhões de jovens, dentre os quais, 58,3 % homens e 41,7 % mulheres. Considerando-se cor ou raça, 27,3 % eram brancos e 71,7 % pretos ou pardos” (IBGE, 2019, p. 10).

A FIGURA 3 abaixo ilustra como são conduzidas as iniciativas de acesso aberto no Brasil para ampliar a abertura do conhecimento científico.

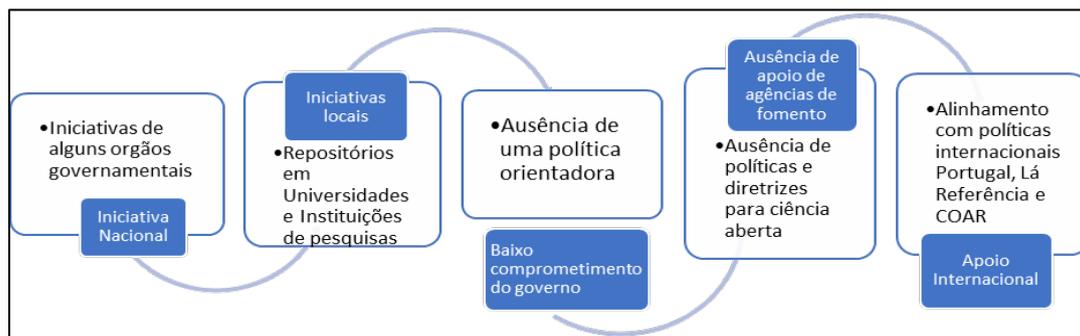


FIGURA 3. Estratégia políticas de acesso aberto no Brasil

Fonte: Elaborado pelas autoras

Assim como os Estados Unidos, o Brasil iniciou as ações buscando aprovação de uma política central, não conseguindo sucesso, investiu em ações em universidades e instituições de pesquisa. Buscou também alinhamento com as políticas internacionais e colabora com Portugal desde 2009 por meio do memorando de entendimento celebrado entre o Ministério da Ciência, Tecnologia e Ensino Superior de Portugal e o Ministério da Ciência, Tecnologia e Inovação do Brasil. Possui acordo de cooperação com a Rede Federada de Repositórios Institucionais de Publicações Científicas da América Latina (LA Referencia) e a Confederação de Repositórios de Acesso Aberto (COAR). Em comparação com os Estados Unidos e Portugal na condução das políticas, observa-se que a falta de apoio dos dois atores principais para estimular o acesso aberto, o governo e as agências de fomento. A ausência do apoio desses atores políticos dificulta o processo de institucionalização das iniciativas existentes. A não priorização do processo de consolidação de uma

política governamental dirigida a todos os atores sociais envolvidos e sem apoio financeiro das agências de fomento à pesquisa reduzem significativamente as chances e o fortalecimento do acesso e ciência aberta no país.

Uma síntese da condução das iniciativas e políticas realizadas pelos três países em estudo é apresentada nos Quadros 1, 2, 3 e 4 a seguir:

INICIATIVAS	ESTADOS UNIDOS	PORTUGAL	BRASIL
Comprometimento do governo em ações de ciência aberta	<p>Representação política de lideranças de acesso aberto nas duas casas de governo.</p> <p>Aprovação de Lei de acesso público, políticas e diretrizes de acesso aberto (ver Quadro 2).</p> <p>Acesso à informação tornou-se prioridade no Governo de Barack Obama.</p> <p>Planos bienais com ampla abertura de dados governamentais e de pesquisa.</p> <p>Investimento em educação, ciência, tecnologia e inovação</p>	<p>As ações de abertura são trabalhadas junto ao Ministério da Ciência, Tecnologia e Ensino Superior (MCTES).</p> <p>Aprovação de legislações direcionadas ao acesso aberto e à ciência aberta.</p> <p>Alinhamento com as políticas de ciência aberta da União Europeia.</p>	<p>As ações de ciência aberta são lideradas pelo IBICT e alguns órgãos governamentais, Universidades e Instituições de pesquisas.</p> <p>As tentativas de aprovação de três projetos de Leis relativas ao acesso aberto foram frustradas (ver Quadro 2).</p> <p>Alinhamento com políticas internacionais.</p>

QUADRO 1. Comprometimento governamental em ações de ciência aberta

Fonte: Elaboração das autoras

Observa-se que, nos Estados Unidos, as ações de ciência aberta são articuladas por líderes governamentais –tanto no Congresso quanto no Senado–, essa estratégia tem contribuído para facilitar a aprovação de projetos de leis nas duas casas do governo. As iniciativas de abertura tiveram destaque a partir de 2009, com a administração do governo Obama. O acesso à informação tornou-se prioridade nos planos de ações governamentais. Como

membro da OGP a partir de 2011, lançou planos bienais com ampla abertura de dados governamentais e de pesquisa. Com altos investimentos em educação, ciência tecnologia e inovação, o presidente assumiu compromissos específicos para promover a transparência pública, promover engajamento cívico, combater a corrupção e ampliar o acesso aos resultados de pesquisa.

No contexto português, as ações de abertura são trabalhadas junto ao Ministério da Ciência, Tecnologia e Ensino Superior (MCTES). A promoção do conhecimento para todos tornou-se um pilar do programa de governo. Em 2016, o governo português publicou a Resolução nº 21 do Conselho de Ministros, na qual são definidos os princípios orientadores e práticas de ciência aberta. Soma-se a essas ações o alinhamento do país com as políticas de acesso aberto da União Europeia.

No Brasil, as ações de ciência aberta são lideradas pelo IBICT e por alguns órgãos governamentais, universidades e instituições de pesquisas. As tentativas do órgão de aprovação de projetos de leis foram frustradas, mas estrategicamente o órgão tem buscado parcerias com outros países e com redes internacionais de repositórios, como La Referencia e COAR.

INICIATIVAS	ESTADOS UNIDOS	PORTUGAL	BRASIL
<p>Legislação Decretos, despachos, resoluções, projetos de leis.</p>	<p>2008 - Primeira política de acesso público do NIH.</p> <p>2017 - Lei Americana de Inovação e Competitividade (AICA)/Lei de <i>Crowdsourcing</i> e Ciência Cidadã</p> <p>2019 - Projeto de Lei do governo bem informado. Proíbe cláusulas de não divulgação em contratos de agências federais com os editores de revistas e fornece uma via para o governo coletar e compartilhar sistematicamente informações sobre o preço das assinaturas de periódicos pagas pelas agências.</p> <p>2019 - Lei de Dados do Governo codifica uma política de dados abertos por padrão” para todos os dados do governo e exige que as agências federais publiquem dados em formatos abertos e legíveis por máquina e usem licenças abertas.</p> <p>2019 - Projeto de Lei Livro Didático reduz o custo dos livros nas faculdades e universidades expandindo o uso de livros abertos (e outros recursos educacionais abertos).</p>	<p>2013 - Decreto de Lei nº 115/2013 (Prevê o depósito obrigatório de uma cópia de teses, dissertações em um repositório integrante da rede RCAAP.</p> <p>2016 - Resolução do Conselho de Ministros nº 21. Define os princípios da implementação da política de Ciência Aberta em Portugal.</p> <p>2016 - Despacho nº 15.389/2016 cria o Grupo de Trabalho (GT-PNCA) para a Política Nacional de Ciência Aberta.</p> <p>2016 - Resolução do Conselho de Ministros nº 21 inclui a obrigatoriedade do Acesso Aberto aos dados científicos financiados por fundos públicos.</p> <p>2017 - Política Nacional de Ciência Aberta - Publicação dos Princípios Orientadores para a Ciência Aberta .</p> <p>2017 - Resolução do Conselho de Ministros cria grupo de trabalho interministerial para elaborar estratégias para Política Nacional de Ciência Aberta.</p>	<p>2007 - Projeto de Lei nº 1.120/2007 dispõe sobre o processo de disseminação da produção técnico-científica pelas instituições de ensino superior no Brasil (arquivado em 2012).</p> <p>2011 - Projeto de Lei do Senado nº 387/2011 obriga as instituições de educação superior de caráter público, bem como as unidades de pesquisa, a construir repositórios institucionais de acesso livre. (Arquivado em 2018).</p> <p>O Projeto de Lei nº 3.702 dispõe sobre a existência de repositórios digitais nas instituições federais de educação superior e de pesquisas (Rejeitado).</p> <p>2018 -Portaria nº 451, de 16 de maio de 2018, define critérios e procedimentos para a produção, recepção, avaliação e distribuição de recursos educacionais abertos.</p>

QUADRO 2. Aparatos legislativos de acesso à informação científica e dados de pesquisa

Fonte: elaboração das autoras

Conforme observado no QUADRO 2, nos Estados Unidos, as ações de ciência aberta são orientadas principalmente pela política de acesso público do NIH – que exige que pesquisadores financiados pelo órgão enviem ao PubMed Central da Biblioteca Nacional de Medicina uma versão eletrônica dos manuscritos revisados por pares depois da aceitação para publicação– e por legislações dirigidas às agências de financiamento à pesquisa. Conforme a liderança entrevistada, “a lei serve de base para avançar e expandir a política para cobrir as pesquisa científica em todas as disciplinas. Ainda não temos uma política de acesso aberto, temos o que chamamos de politica de acesso público, porque há um ano de embargo permitido, por isso defendemos agora o embargo zero e uma política completa de acesso aberto nos Estados Unidos” (Entrevistado E, 2019).

Em Portugal, as legislações definem princípios, orientações, infraestrutura e estratégias para trabalhar a aprovação de uma política nacional de ciência aberta. Conforme o entrevistado, “para ter uma política nacional não basta ter uma política central e depois não existir nada. É necessário infraestrutura, suporte dos diferentes atores, das instituições e das pessoas, é depois que essa política se aplica” (Entrevistado P, 2018).

No Brasil, os projetos de leis que poderiam proporcionar melhor infraestrutura e orientações para universidades e instituições de pesquisas não foram aprovados. Identificou-se uma portaria que define critérios e procedimentos para a produção, recepção, avaliação e distribuição de recursos educacionais abertos, no entanto, não foram identificados aparatos legislativos direcionados às agências de fomento. De acordo com o entrevistado, “o limitante das ações políticas é a ausência de lei. Houveram duas tentativas frustradas de criação da Lei de acesso à publicação científica que tramitou no governo por anos e não foram aprovadas (Entrevistado B, 2019).

INICIATIVAS	ESTADOS UNIDOS	PORTUGAL	BRASIL
Política mandatória para agências de fomento à pesquisa	Lei FASTR instrui as agências federais no desenvolvimento de políticas de acesso público relativas a pesquisas financiadas com fundos públicos e administrados pelo governo  Plano de Ação Nacional do governo Obama orientou as agências sobre práticas, treinamento e incentivo para ciência cidadã e <i>crowdsourcing</i>	Política de acesso Aberto FCT. Obrigatoriedade de disponibilização em Acesso Aberto às publicações de pesquisas financiadas  Alinhamento com as políticas de outras agências de financiamento da União Europeia	A pesquisa não identificou política governamental direcionada às agências de financiamento  Política de acesso aberto da FAPESP

QUADRO 3. Política mandatória para agências de fomento à pesquisa

Fonte: Elaboração das autoras

Os Estados Unidos têm política mandatória para as agências de financiamento. As agências federais são orientadas a ampliar o uso de métodos de inovação aberta e estabelecer políticas de incentivo à ciência cidadã e crowdsourcing. Em Portugal, a maior agência de financiamento implementou política de obrigatoriedade de disponibilização em acesso aberto às publicações de pesquisa financiadas com recursos públicos; o país está alinhando com as suas políticas locais às de outras agências de fomento da União Europeia. No Brasil, não foi possível identificar políticas direcionadas às agências financiadoras de pesquisa. Nesse sentido, das 29 principais agências que financiam pesquisas no país (três federais: CAPES, CNPq e FUNEP; e 26 estaduais: FAPs), até 2020, somente a FAPESP havia instituído uma política de acesso aberto.

INICIATIVAS POLITICAS	ESTADOS UNIDOS	PORTUGAL	BRASIL
<p>Engajamento de instituições ligadas à informação e à cultura no movimento de acesso aberto</p>	<p>Forte mobilização social pelo acesso aberto à literatura científica</p> <p>Mobilização das bibliotecas acadêmicas e de pesquisa, faculdades, universidades e grupos de pesquisa e organizações de todo o país</p> <p>Carta dirigida ao presidente Donald Trump, assinada por organizações, cientistas, pesquisadores e associações, solicitando a atualização da política de acesso público e a eliminação do período de embargo de 12 meses</p>	<p>Iniciativas são lideradas pelo Centro de Documentação da Universidade do Minho, pelas instituições de ensino superior e pelo grupo de trabalho sobre o acesso aberto</p> <p>A forte atuação do Conselho de Reitores mobiliza o apoio do governo e fortalece as iniciativas de políticas das universidades</p> <p>Orgãos ligados à educação, informação, tecnologia e cultura têm representação no Comitê Interministerial de ciência aberta</p>	<p>O apoio do IBICT, somado aos esforços de muitas universidades e instituições de pesquisa, tem contribuído para ampliar o número de repositórios digitais no país</p> <p>Mobilização das instituições que compõem a Rede Brasileira de Repositório Institucional da Produção Científica em Acesso Aberto</p> <p>A pesquisa não identificou grandes mobilizações (em conjunto) de organizações como bibliotecas, museus, arquivos, representações de professores, alunos e outras associações ligadas à informação e à cultura no movimento pela ciência aberta</p>

QUADRO 4. Engajamento de instituições ligadas à informação e à cultura no movimento de acesso aberto

Fonte: Elaboração das autoras

A intensa mobilização da sociedade, que reivindica o acesso público gratuito à literatura, foi fundamental para a sensibilização do governo estadunidense sobre a importância do acesso aberto. A mobilização conta com a participação de 210 bibliotecas acadêmicas e de pesquisa, faculdades, universidades e organizações de todo o país, entre as quais 66 organizações nacionais de

defesa de pacientes com doenças, 45 associações nacionais estudantis e mais de 149.000 representantes do Grupo de Pesquisa de Interesse Público e dos estudantes. Sobre este processo de sensibilização, a entrevistada relata que a principal barreira foi convencer os formuladores de política sobre o que são periódicos científicos e artigos de periódicos e porque eles são importantes “às vezes eles apenas olhavam para nós e falavam: o que é isso mesmo? Então nós levamos alguns anos para convencê-los”.

Em Portugal, A forte atuação do Conselho de Reitores mobiliza o apoio do governo e fortalece as iniciativas de políticas das universidades. Além disso, a liderança do Centro de Documentação da Universidade do Minho mobiliza ações com os órgãos ligados à educação informação, tecnologia e cultura. Órgãos ligados à educação informação, tecnologia e cultura têm representação no Comitê Interministerial de ciência aberta. Conforme o entrevistado “o trabalho de conscientização é realizado na base com pessoas executoras de políticas. Nesse trabalho tinha pessoas da tecnologia de informação e um conjunto de pessoas de organizações em subgrupos de trabalho para tentar definir para quem a política vai ser aplicada, procuramos não fazer uma política só de cima para baixo.” (Entrevistado P, 2018).

No Brasil, o apoio do IBICT, somado aos esforços de muitas universidades e instituições de pesquisa, tem contribuído para ampliar o número de repositórios digitais no país. Observou-se o engajamento das instituições que compõem a Rede Brasileira de Repositório Institucional da Produção Científica em Acesso Aberto. “antigamente o tema ciência aberta era falado por poucas instituições, agora tem muita instituição envolvida e tratando do assunto, entre as quais Fiocruz, Embrapa, Capes, CNPq, SciELO, Abec, confap” (Entrevistado B, 2019). Entretanto, não foram identificadas mobilizações coletivas voltadas à promoção do acesso aberto aos dados das pesquisas.

Os conteúdos sintetizados nos Quadros 1, 2,3 e 4 não esgotam, mas traçam indicadores que auxiliam na identificação de algumas das ações do movimento político do acesso aberto nos três países estudados e evidenciam o empenho e o comprometimento e o percurso adotado por estas nações na abertura do conhecimento científico.

## 5. Conclusão

No estudo realizado observou-se que as ações políticas são determinantes para ampliar e consolidar a ciência aberta nos estados nacionais, contudo não se circunscrevem apenas aos aparatos legislativos. A ciência aberta requer uma política pública forte e direcionada para as ações de abertura que inclua o engajamento de atores políticos, sociais e econômicos. Na elaboração desta política é necessário considerar especialmente a realidade local, as vulnerabilidades e assimetrias existentes de modo que as medidas de mitigação de danos possam ajudar efetivamente na redução das barreiras sociais, minimizar as diferenças e as desigualdades históricas principalmente com relação ao acesso à educação, para que o cidadão se reconheça socialmente como sujeito de direitos e deveres e compreenda a centralidade da informação e do conhecimento científico no exercício da cidadania.

Considerando-se o estágio de desenvolvimento das políticas de acesso ao conhecimento científico nos Estados Unidos, Portugal e Brasil, depreende-se que a efetivação das políticas que orientam, incentivam e regulamentam a ciência aberta requer um modelo de governança em rede com o estímulo ao engajamento de distintos atores políticos e sociais, sendo o governo e as agências de fomento imprescindíveis. Tais iniciativas, além de compreenderem o acesso imediato aos resultados da fortuna científica produzida em território nacional, precisam considerar a adoção sistemática de medidas estruturais que incorporem políticas educacionais, de inclusão digital e de acesso à informação e à cultura como aspectos prioritários ao exercício da justiça epistêmica enquanto uma dimensão fundamental da cidadania.

## Bibliografía

- ALBAGLI, S. (2015). Ciência aberta em questão In: Albagli S. Maciel, M. L & ABDO, A. H. (Org.). Ciência aberta, questões abertas. (p. 9-25). Brasília: IBICT; Rio de Janeiro: UNIRIO.
- BARDIN, L. (2011). Análise de conteúdo. São Paulo: Edições, 70.
- BRAMAN, S. (2011). Defining information policy. *Journal of Information Policy*, 1: 1-5.  
[http://people.tamu.edu/~Braman/bramanpdfs/38\\_defininginfopolicy.pdf](http://people.tamu.edu/~Braman/bramanpdfs/38_defininginfopolicy.pdf)
- CARUSO, J.& NICOL, A.; ARCHAMBAULT, E. (2014). Comparative analysis of the strengths and weaknesses of existing open access strategies. RTD-B6-PP-2011-2: Study to develop a set of indicators to measure open access. Montreal: Science-Metrix Inc.  
[https://www.science-metrix.com/sites/default/files/sciencemetrix/publications/d\\_2.3\\_sm\\_ec\\_dg-rtd\\_oa\\_comparative\\_analysis\\_update\\_v04p\\_0.pdf](https://www.science-metrix.com/sites/default/files/sciencemetrix/publications/d_2.3_sm_ec_dg-rtd_oa_comparative_analysis_update_v04p_0.pdf)
- ESTADOS UNIDOS (2013). The Open Government Partnership. Second Open Government National Action Plan for the United States of America. Feb.
- ESTADOS UNIDOS (2015). The Open Government Partnership. Third Open Government National Action Plan for the United States of America. Oct.
- ESTADOS UNIDOS (2008). National Institutes of Health (NIH). NIH Office of Science Policy, Recuperado de <https://osp.od.nih.gov/>.
- ESTADOS UNIDOS (2011). The Open Government Partnership. First Open Government National Action Plan for the United States of America.
- ESTADOS UNIDOS (2019). H.R.5049. WISE Government Act. Well-Informed, Scientific, & Efficient Government Act of 2019 or the WISE Government Act. Introduced in House
- ESTADOS UNIDOS (2019). The Open Government Partnership. Fourth Open Government National Action Plan for the United States of America. Feb.
- ESTADOS UNIDOS. (2017). American Innovation and Competitiveness Act (AICA). Section 402 Crowdsourcing and Citizen Science Act. <https://www.congress.gov/bill/114th-congress/senate-bill/3084/text/enr>
- ESTADOS UNIDOS. (1966). Freedom of Information Act – FOIA.  
<https://www.foia.gov/about.html>

- FECHER, B; FRIESIKE, S. (2013). Open Science: One Term, Five Schools of Thought. May 30, RatSWD\_WP\_ 218.  
[http://ssrn.com/abstract=2272036srn.com/sol3/papers.cfm?abstract\\_id=2272036](http://ssrn.com/abstract=2272036srn.com/sol3/papers.cfm?abstract_id=2272036)
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE) (2018). PNAD Contínua: Educação.  
[https://biblioteca.ibge.gov.br/visualizacao/livros/liv101657\\_informativo.pdf](https://biblioteca.ibge.gov.br/visualizacao/livros/liv101657_informativo.pdf)
- JORANSON, K. (2008). Indigenous knowledge and the knowledge commons. *Journal International Information & Library Review*, 40(1): 64-72.
- MAZZUCATO, M. (2018). Mission-oriented research & innovation in the European Union: a problem-solving approach to fuel innovation-led growth. Luxembourg: Publications Office of the European Union.  
[https://www.ucl.ac.uk/bartlett/publicpurpose/sites/public-purpose/files/mission-oriented\\_ri\\_in\\_the\\_eu\\_mazzucato\\_2018.pdf](https://www.ucl.ac.uk/bartlett/publicpurpose/sites/public-purpose/files/mission-oriented_ri_in_the_eu_mazzucato_2018.pdf)
- MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR (2016). Ciência Aberta: o conhecimento é de todos para todos.  
<https://www.portugal.gov.pt/media/18506199/20160210-mctes-ciencia-aberta.pdf>
- PATTERSON, T. E. (2018). We The People: an introduction to American government. (13. Ed.). [S. I.]: MacGraw Hill Education.
- SILVA, P. F. P. (2017). As políticas de open data em Portugal: análise da sua implementação e impacto. (Dissertação de Mestrado em Ciência da Informação, Departamento de Filosofia, Comunicação e Informação da Faculdade de Letras da Universidade de Coimbra, Coimbra, Portugal).
- SPARC. (2020) Open education. <https://sparcopen.org/open-education/>
- SWAN, A. (2012). Diretrizes para as políticas de desenvolvimento do acesso aberto. Paris: Unesco. <http://unesdoc.unesco.org/images/0021/002158/215863e.pdf>
- THE ROYAL SOCIETY. (2012). Science as an Open Enterprise. The Royal Society Policy Centre Report. <https://royalsociety.org/topics-policy/projects/sciencepublicenterprise/report/>
- UNESCO Recommendation on Open Science. (2021). Organização das Nações Unidas para a Educação, a Ciência e a Cultura. Unesco.  
[https://unesdoc.unesco.org/ark:/48223/pf0000379949\\_por](https://unesdoc.unesco.org/ark:/48223/pf0000379949_por)
- YIN, R. K. (2015). Estudo de caso: planejamento e métodos. (5. ed.) Tradução de Cristian Matheus Herrera. São Paulo: Bookman.

# Transformación digital de un archivo histórico mediante el uso de un software de código abierto: el caso del Archivo Histórico de la Real Caja de Zacatecas

**SARA MARÍA DEL PATROCINIO RODRÍGUEZ PALACIOS**

Tecnológico de Monterrey (México)

[sara.rodriguez@tec.mx](mailto:sara.rodriguez@tec.mx)

## RESUMEN

Las entidades y agentes culturales tienen como misión administrar y gestionar sus bienes y servicios, estableciendo para ellos los procedimientos, normas y herramientas que garanticen el logro de sus objetivos. En su compromiso por preservar, dar visibilidad y acceso a sus contenidos, la Dirección Nacional de Patrimonio Cultural del Tecnológico de Monterrey gestionó la digitalización del Archivo Histórico de la Real Caja de Zacatecas, 1576-1936, mismo que custodia, conserva, difunde y tiene en comodato su Campus Zacatecas. En esta ponencia se compartirán algunos aspectos que se contemplaron para garantizar el éxito de este proceso: organización de los contenidos para reflejarlos en la herramienta tecnológica de gestión del Repositorio Institucional del Tecnológico de Monterrey (RITEC) para reflejar el cuadro de clasificación archivística; ajustes a la política de desarrollo de colecciones, definición de niveles de acceso, aprovisionar las capacidades de almacenamiento de información o de transferencia de datos; apego a un estándar de metadatos que permitiera registrar las características particulares; los procesos de preservación, la curaduría de metadatos y la importante colaboración de un gran equipo

interdepartamental e interdisciplinario. Se comparten los resultados y se propone una innovación tecnológica como estrategia de mejora para la gestión de los recursos digitalizados.

#### **ABSTRACT**

*The mission of cultural entities and agents is to administer and manage their goods and services by establishing procedures, norms and tools that guarantee the achievement of their objectives. In its commitment to preserve, give visibility and access to its contents, the National Direction of Cultural Heritage of the Tecnológico de Monterrey managed the digitization of the Historical Archive of the Real Caja de Zacatecas, 1576-1936, which is kept, conserved, disseminated and on loan at its Zacatecas Campus. In this paper we will share some aspects that were contemplated to guarantee the success of this process: organization of the contents to reflect them in the technological management tool of the Institutional Repository of the Tecnológico de Monterrey (RITEC) in order to reflect the archival classification chart; adjustments to the collection development policy, definition of access levels, provisioning of information storage or data transfer capacities; adherence to a metadata standard that allowed the recording of particular characteristics; preservation processes, metadata curation and the important collaboration of a large interdepartmental and interdisciplinary team. The results are shared and a technological innovation is proposed as an improvement strategy for the management of digitized resources.*

#### **PALABRAS CLAVE**

Acceso abierto; patrimonio cultural; transformación digital; repositorios de acceso abierto.

#### **KEYWORDS**

*Open access; cultural heritage; digital transformation; open access repositories.*

## 1. Introducción

El Archivo Histórico de la Real Caja de Zacatecas, 1576-1936 forma parte de la Colección Sociedad de Amigos de Zacatecas, A. C., es uno de los principales archivos históricos zacatecanos. Como menciona MAURICIO-ESCALANTE (2021), durante décadas este acervo documental estuvo en la Biblioteca Clements de la Universidad de Michigan en Ann Arbor y regresó a Zacatecas en 1992, gracias a las gestiones del Lic. Manuel Fernando Sescosse Varela. En este acervo documental se encuentra la administración y recaudación de impuestos que se realizaron desde el último cuarto del siglo XVI hasta principios del siglo XX, en la ciudad de Zacatecas. La minería, el comercio, las misiones religiosas al Norte de México, la milicia que se estableció en la ciudad durante el siglo XIX, el reparto de azogue y sal de las Salinas de Santa María y Del Peñol Blanco, el Colegio Seminario de San Luis Gonzaga, entre otros.

## 2. Desarrollo

### 2.1. Marco teórico

Los archivos históricos resguardan información social, política y económica de las sociedades. BABINES LÓPEZ hace hincapié en los valores secundarios: evidenciales, testimoniales e informativos.

Los evidenciales son los que muestran derechos y obligaciones en una sociedad; en este rubro podemos englobar a las leyes, reglamentos, decretos, etc., considerados base de la historia institucional. Los testimoniales nos muestran la evolución de las instituciones, se asigna este valor a los documentos generados por funciones sustantivas de la institución. Y por último, los informativos, documentos que sirven de instrumento de información para el usuario, la entidad productora y como fuente para la investigación política, social, demográfica, cultural, religiosa, etc. (BABINES LÓPEZ, 2022)

Los contenidos de los archivos históricos son, o deberían ser, para la sociedad. Sin embargo, temas como la preservación física (son documentos que requieren ciertas condiciones de luz, humedad, etc.), la ubicación geográfica (un investigador puede tener que hacer largos viajes para poder analizar los contenidos de los archivos), las gestiones requeridas para consultas (cartas, justificaciones, cuadrar agendas) y otros que se puedan acumular se convierten en una barrera difícil de sortear.

La transformación digital parece ser la forma como se puede garantizar el acceso y visibilidad más inclusivos. En palabras de TORRES (2020), esta transformación digital requiere tres procesos sucesivos:

- digitización (conversión de la información analógica a un formato digital) de los recursos documentales;
- digitalización (proceso de cambio) que implica garantizar el establecimiento de procesos, normas, técnicas y herramientas que garanticen la gestión, acceso y visibilidad de los recursos, y
- transformación digital (efecto o consecuencia de la digitalización de los procesos de negocios) que implica un cambio cultural, el cual involucra la participación de los contenidos en la vida digital de las organizaciones y sociedades.

## **2.2. Descripción y proceso de implementación de la innovación**

La transformación digital del Archivo Histórico de la Real Caja de Zacatecas, 1576-1936 debió tener en cuenta muchos aspectos; se hará mención de aquellos más relacionados a la gestión digital.

Se seleccionó un software de código abierto que provee herramientas para la administración de colecciones digitales, DSpace en su versión 6.3, ya que es el medio utilizado por el Repositorio Institucional del Tecnológico de Monterrey (RITEC), sitio donde se decidió agregar estos recursos por ya existir una Comunidad de Patrimonio Cultural. El RITEC tiene como propósito ser la ventana única donde el Tecnológico de Monterrey comparte todos los recursos científicos, académicos y administrativos que genera, así como el patrimonio

documental y artístico que se encuentra bajo su resguardo y que está disponible en Acceso Abierto.

El RITEC se organiza utilizando la estructura ofrecida por DSpace, donde las comunidades son el nivel más alto de la jerarquía de contenidos. Pueden estar formadas por subcomunidades y por colecciones, estas últimas son las que contienen los ítems directamente, tal como se muestra en el GRÁFICO 1.

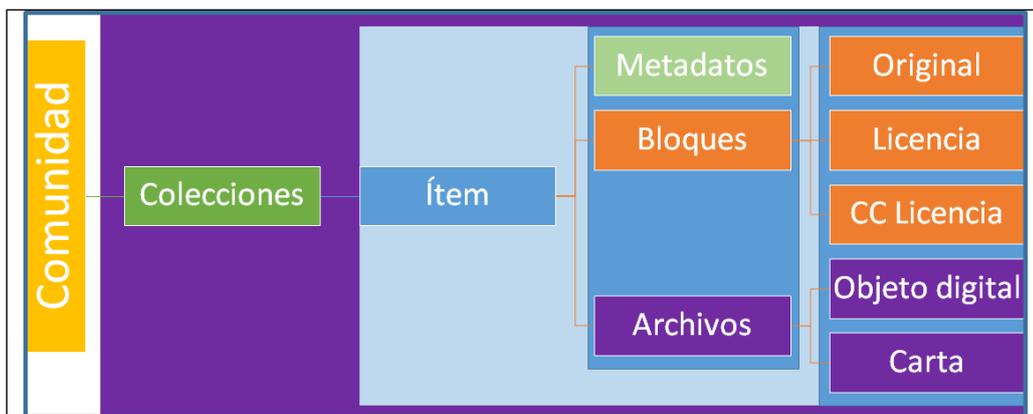


GRÁFICO 1. Estructura de DSpace 6.x

Como puede observarse, una Comunidad es el nivel más alto de la jerarquía de contenidos de DSpace. Pueden estar formadas por Subcomunidades y por Colecciones, estas últimas son las que contienen los ítems directamente.

Se realizó un análisis profundo del cuadro de clasificación archivística para ajustar sus necesidades a la estructura que ofrece DSpace, por lo cual se pudo realizar el ajuste de manera exitosa, quedando como se muestra en el ejemplo del GRÁFICO 2:

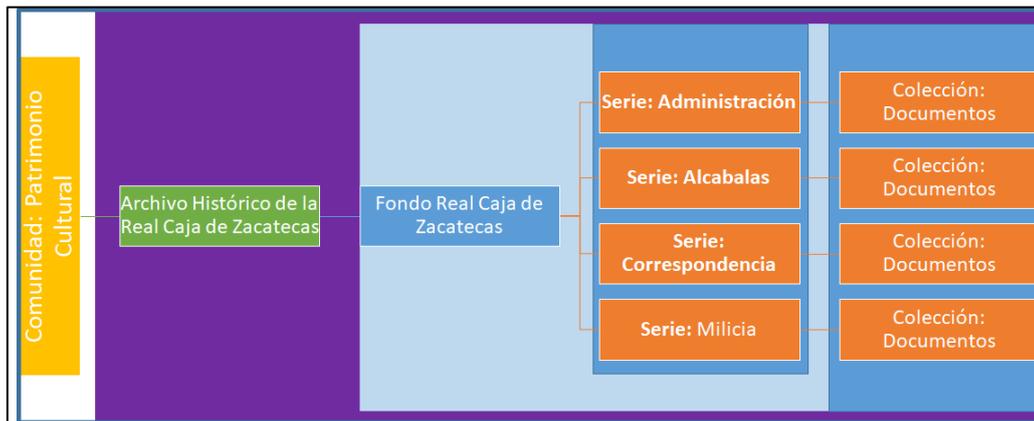


GRÁFICO 2. Estructura del Archivo Histórico de la Real Caja de Zacatecas, Fondo, Series y Colecciones

La política del desarrollo de colecciones del RITEC contemplaba inicialmente a todas las comunidades en general; sin embargo, hubo que hacer una actualización para ajustar las condiciones particulares que aplican a los materiales del archivo histórico, entre otros:

- formatos de ficheros
  - a) divulgación en PDF/A con reconocimiento óptico de caracteres (OCR)<sup>2</sup>
  - b) de preservación en TIFF
- nivel de acceso
  - a) formatos de divulgación en acceso abierto con Licencia Creative Commons específica por colección
  - b) formatos de preservación con acceso restringido, solo se permite su lectura a los administradores de las colecciones

Una vez digitalizados los recursos, se realizaron cargas masivas al RITEC mediante la herramienta Batch Metadata Editing, que permite realizar importaciones de ítems completos, que quedan depositados en el Repositorio en acceso abierto. Se siguió el proceso señalado en el GRÁFICO 3.

<sup>2</sup> El estándar PDF/A: PDF para la preservación a largo plazo. <https://es.scribd.com/document/240311146/m2u2-anexo09-estandar-pdf-a-pdf>

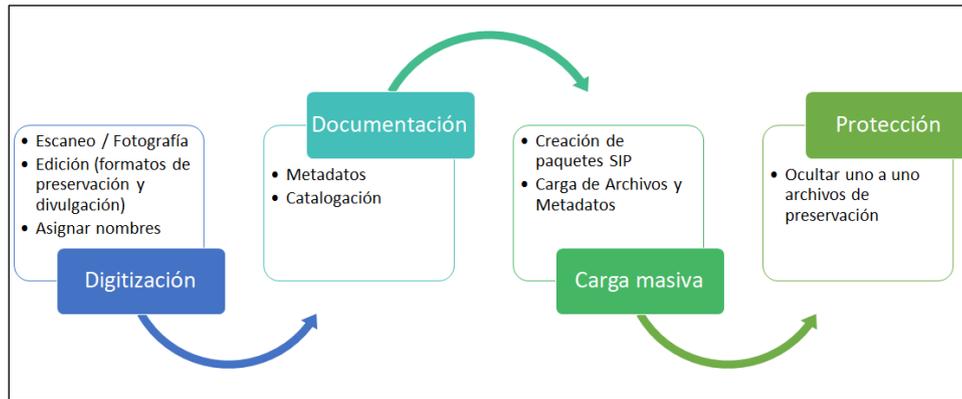


GRÁFICO 3. Flujo de depósito en el RITEC de recursos digitalizados

El sistema DSpace permite regular el nivel de acceso desde cualquiera de sus niveles, y siempre las modificaciones de un nivel afectan a todos los subniveles que contiene; por lo cual nos encontramos con un reto muy importante, ya que se tuvo que comenzar a hacer la protección de acceso solamente a los objetos digitales de preservación, esto implicó una mayor inversión económica y de tiempo; en el caso de ítems con archivos muy numerosos se tuvo que programar la restricción de acceso uno a uno. Como se muestra en el GRÁFICO 4, para el caso de una foja el trabajo que esto implica puede parecer rápido, sin embargo, cuando hablamos de cargas masivas y de expedientes numerosos surgen al menos tres áreas de oportunidad:

1. Seguridad (ventana de tiempo) los archivos de preservación que deben estar protegidos se encuentran expuestos en tanto se termina de realizar el cambio de nivel de acceso.
2. Tiempo invertido en protección. Es muy significativo.
3. Operación (no depender de horarios de carga). Como debe asegurarse la protección de todos los archivos de preservación tan rápido como sea posible, se tuvo que restringir el horario en que se realizaban las cargas masivas, para que el equipo colaborador tuviera el tiempo suficiente dentro de sus horas laborales para realizar la protección uno a uno.



GRÁFICO 4. Áreas de oportunidad para protección de archivos de preservación

A raíz de esta situación, se gestionó ante el equipo de administración tecnológica el desarrollo de un módulo que permitiera realizar estas cargas masivas con ficheros protegidos donde se pudieran asignar los privilegios de lectura de cada fichero a grupos de usuarios definidos. Este desarrollo permitió la eficiencia en tres rubros, como se muestra en el GRÁFICO 5, un menor costo ya que se utilizaron menos horas de trabajo del equipo especializado para realizar esta tarea; ahorro en tiempo ya que el añadir las características de nivel de acceso en el repositorio es inmediato en cuanto se realiza la carga masiva y, finalmente, una mejora muy significativa en cuanto al nivel de seguridad, pues ningún recurso queda expuesto desde el momento en que es depositado.

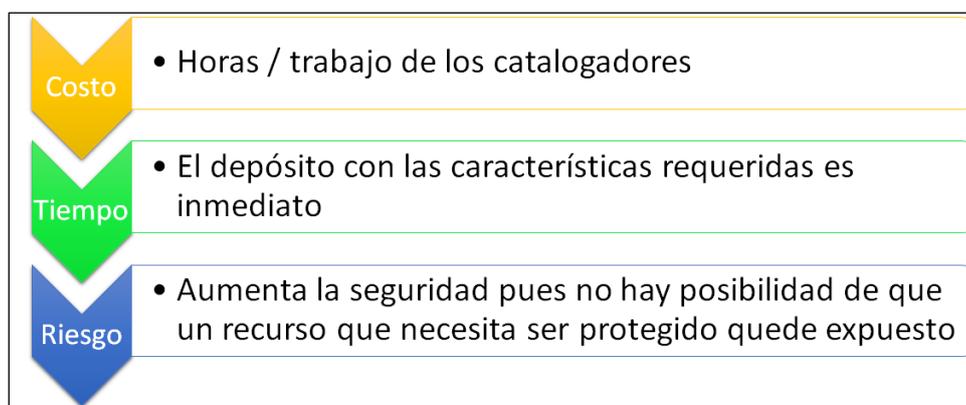


GRÁFICO 5. Eficiencia del desarrollo de cargas masivas protegidas en el repositorio

Fueron requeridas cuatro semanas para el desarrollo de la herramienta que permitiera cubrir los siguientes requerimientos:

1. Generar estructuras de importación, que pueda proteger los ítems e indicar al sistema que suben en oculto.
2. El habilitador debe indicar, con una clave o con un metadato adicional en la hoja de cálculo donde se registran los metadatos, si un fichero asociado a un ítem tendrá permisos restringidos.
3. Que funcione con interfaz web preferentemente, o en su defecto, en modo consola (indicando los pasos a seguir para que una persona de la institución pueda hacer este trabajo en modo consola, e indicando también los comandos Linux de base).
4. Capacidad para seleccionar el grupo específico que tendrá acceso a los ficheros, básica y prioritariamente el Administrator y en segundo lugar algún otro que se pueda seleccionar dependiendo de las necesidades específicas de diseminación de las colecciones.

Tras esta actividad se implementó el flujo para cargas masivas protegidas, que tal como se muestra en el Gráfico 6 contempla los siguientes puntos:

#### *Pasos previos*

1. Se genera una hoja de cálculo en Excel donde cada columna se encabeza con la etiqueta Dublin Core de los metadatos a agregar. En una columna adicional se agregan los nombres de archivos y el nombre del Grupo que tiene privilegio de lectura.
2. Utilizando un habilitador desarrollado para esto, se generan carpetas SIP que relacionan los metadatos y sus respectivos archivos. Estos archivos se comprimen en un ZIP.

### Carga masiva desde interfaz web

3. El usuario con privilegios de Administrator ingresa al Repositorio con su cuenta.
4. Selecciona la Herramienta Importación Batch (ZIP)
5. Elige la colección en donde se estarán depositando los recursos.
6. Selecciona el archivo ZIP generado.
7. Da clic en el botón [Subir un ZIP SimpleArchiveFormat]
8. DSpace genera un reporte con los recursos agregados.

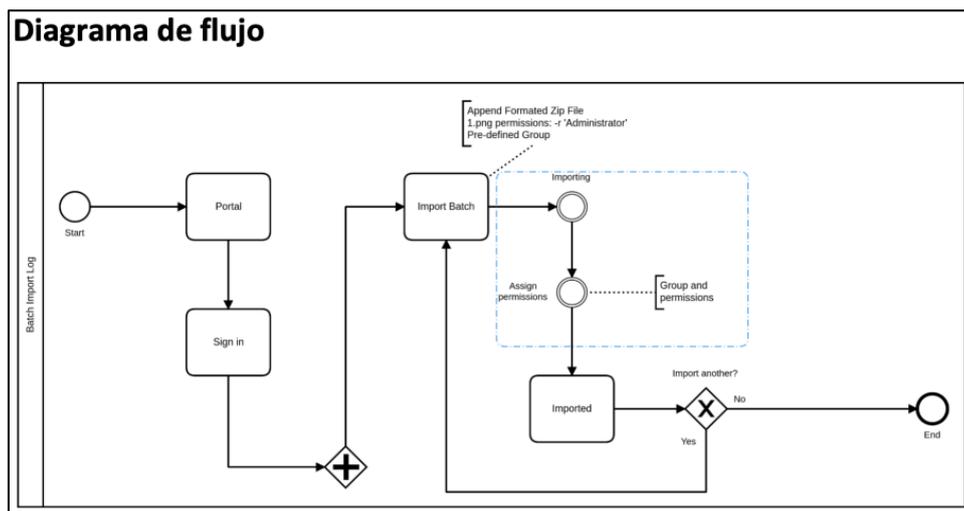


GRÁFICO 6. Diagrama de flujo de cargas masivas protegidas en el repositorio

Los metadatos son muy importantes para comprender qué contienen los recursos, además de aportar información indispensable sobre su estructura y son utilizados para su gestión administrativa. El cumplir con estándares internacionales garantiza la buena gestión, proporciona vínculos entre los recursos y su contexto de creación, permite la interoperabilidad para reconocer, procesar y usar en otros entornos, se cumple con requisitos legales de uso y distribución, reduce el riesgo de acceso no autorizado y facilitan los procesos de conversión, migración y conservación a largo plazo. Por ejemplo, en el caso de metadatos de fechas pues no aplica el estándar de uso exclusivo del formato ISO 8601 que dicta los cuatro dígitos del año, guion, dos dígitos para mes, guion, dos dígitos para día: AAAA-MM-DD. Las colecciones especiales y

en particular este archivo histórico requiere una entrada de texto abierto en fecha de creación para poder hacer anotaciones correspondientes a fecha cierta o aproximada, tales como: década probable, década segura, siglo probable, siglo seguro, fecha probable, posterior a, anterior a, cerca de o un periodo de tiempo comprendido entre dos fechas.

Aunado al esfuerzo de integrar estándares internacionales se realizó un trabajo colaborativo con el equipo especializado de catalogación para la asignación de metadatos y vocabularios controlados pertinentes, permitiendo que esta labor de curaduría refleje parte de la riqueza y belleza de las características particulares de cada tipo de recurso. Como resultado, se logró la creación de los siguientes documentos:

- Fuentes de consulta del vocabulario controlado para el Repositorio Institucional Tec (RITEC)<sup>3</sup>
- Glosario de términos bibliográficos, archivísticos, fotográficos y museográficos<sup>4</sup>

Se realizó, además, una proyección para determinar con anticipación las capacidades de almacenamiento de información o de transferencia de datos que serían requeridas, se muestra en la TABLA 1.

Tipo de recurso	Cantidad de ítems	Promedio de archivos por ítem		Peso unitario MB		Peso promedio Mb por ítem	Peso TOTAL Mb	Peso TOTAL Tb
		TIFF	PDF/A	TIFF	PDF/A			
Documentos y libros	8,710	20	1	25	150	650	5661500	5.3

TABLA 1. Cálculo de espacio para los recursos digitales del Archivo Histórico de la Real Caja de Zacatecas, 1576-1936 en el RITEC

<sup>3</sup> Fuentes de consulta del vocabulario controlado para el Repositorio Institucional Tec (RITEC): <https://hdl.handle.net/11285/636607>

<sup>4</sup> Glosario de términos bibliográficos, archivísticos, fotográficos y museográficos: <https://hdl.handle.net/11285/636686>

Dado este análisis, hubo que aprovisionar una capacidad de almacenamiento en el RITEC de 5,5 TB mediante Amazon S3 Glacier<sup>5</sup>, el cual ofrece el almacenamiento para los datos de larga duración que requieren una recuperación inmediata y copias de seguridad. La capacidad de almacenamiento requerida fue de 5,4 TB, muy similar a la proyectada. Adicionalmente se tiene una copia de todos los archivos en un disco externo.

### 2.3. Evaluación de resultados

Podemos comenzar a visualizar el gran impacto de la transformación digital de este acervo, simplemente en el periodo enero-abril 2021, el Archivo Histórico de la Real Caja de Zacatecas, 1576-1936 ha tenido 906 visitas y 858 descargas, se distribuyen como se muestra en el GRÁFICO 6.

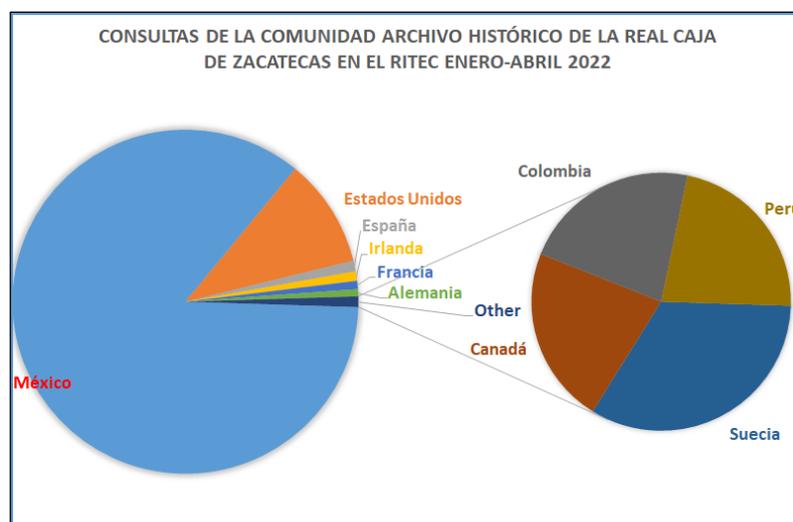


GRÁFICO 6. Visitas a ítems de las colecciones del Archivo Histórico de la Real Caja de Zacatecas en el RITEC en el periodo enero-abril 2022. Fuente:

<https://repositorio.tec.mx/handle/11285/636333>

La gran parte de visitas es de México; sin embargo, podemos observar que se extienden también hasta otros países y continentes. En la medida que se siga dando a conocer este acervo se tiene la expectativa que pueda tener un

<sup>5</sup> Amazon S3 Glacier: <https://aws.amazon.com/es/s3/storage-classes/glacier/?refid=5970b1e9-218b-48cc-9862-f23c151d81b2>



Repositorio Institucional que, en conjunto con un visor de imágenes IIIF, brinde soporte para permitir la activación de la funcionalidad de búsqueda dentro del visor de IIIF, proporcionando búsqueda dentro de las imágenes, navegación a través de los resultados y resaltado en la imagen del texto OCR correspondiente a los términos de búsqueda ingresados. Se está ponderando el depósito de formatos de archivos de preservación que permitirán, además de garantizar la accesibilidad permanente de los objetos digitales, más adelante implementar estas iniciativas.

Adicional a ello, se sugiere fortalecer la preservación de los recursos digitalizados implementando un servicio de alta capacidad de almacenamiento y duplicación de la información que interactúe con el RITEC mediante la sincronización de metadatos, de tal manera que se eficienten los costos de almacenamiento y transferencia de datos.

Es de gran importancia dar visibilidad y acceso a estas colecciones patrimoniales como una gran estrategia para corresponder a los ideales de los Objetivos de Desarrollo Sostenible (ODS)<sup>6</sup>. Los contenidos de un archivo histórico que se pueden volver inaccesibles por temas de preservación física, podrán estar disponibles al mundo entero gracias al uso de la tecnología de manera equitativa e inclusiva, mediante procesos de transformación digital que promuevan oportunidades de aprendizaje permanente para todos.

Cerramos agradeciendo a todo el equipo que ha participado en el flujo de trabajo completo: la organización previa del archivo, la definición del cuadro de clasificación archivística, la revisión y documentación de plantillas de metadatos, la digitización, envío de archivos, recepción, catalogación y documentación, así como la gestión general de las colecciones patrimoniales en el Repositorio Institucional del Tecnológico de Monterrey.

---

<sup>6</sup> Objetivos de Desarrollo Sostenible: <https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>

## Bibliografía

- AVILA BANDA, H. (2020, August 15). Glosario de Términos bibliográficos, Archivísticos, fotográficos y museográficos. Repositorio Institucional del Tecnológico de Monterrey. <https://hdl.handle.net/11285/636686>
- AVILA BANDA, H. (2020, July 23). Fuentes de Consulta del Vocabulario Controlado para el Repositorio institucional TEC (RITEC). Repositorio Institucional del Tecnológico de Monterrey. <https://hdl.handle.net/11285/636607>
- BABINES LÓPEZ, J. (n.d.). La importancia de los documentos históricos. – Archivo General del Estado de Oaxaca. <https://www.oaxaca.gob.mx/ageo/la-importancia-de-los-documentos-historicos/>
- EMAZ, A., & PÉREZ-RUIZ U. (2021). Amazon S3 Glacier. Amazon. <https://aws.amazon.com/es/s3/storage-classes/glacier/?refid=5970b1e9-218b-48cc-9862-f23c151d81b2>
- MAURICIO ESCALANTE, Z. M. (n.d.). Archivo Histórico de la Real Caja de Zacatecas custodiado por el Tecnológico de Monterrey en Campus Zacatecas. AMABPAC. <https://www.amabpac.org.mx/wp/archivo-historico-de-la-real-caja-de-zacatecas-custodiado-por-el-tecnologico-de-monterrey-en-campus-zacatecas/>
- MAURICIO ESCALANTE, Z. M., & MACÍAS CHIU, A. L. (2021, November 23). El problema de la Memoria en colecciones digitales. Preservación digital del patrimonio cultural. Retos, aprendizajes y reflexiones desde el Tecnológico de Monterrey. Patrimonio en un bit. <https://muac.unam.mx/patrimoniobit/en-vivo/el-problema-de-la-memoria-en-colecciones-digitales.html>
- PDF ASSOCIATION. (2022). PDF/a. <https://www.pdfa.org/wp-content/uploads/2011/07/pdfa-flyer-esp.pdf>
- TORRES, J. (2020, July 30). Los retos de la transformación digital para los repositorios y las bibliotecas digitales. figshare. [https://figshare.com/articles/presentation/Dia\\_2\\_Joel\\_Torres\\_Transformacio\\_n\\_digital\\_pptx/12736124](https://figshare.com/articles/presentation/Dia_2_Joel_Torres_Transformacio_n_digital_pptx/12736124)
- UNITED NATIONS. (2020). Objetivos y Metas de Desarrollo Sostenible - Desarrollo Sostenible. United Nations. <https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>

# Comunicación de la ciencia, la tecnología y la innovación: elementos para una conceptualización democratizadora

## **IRENE TRELLES RODRÍGUEZ**

Universidad Católica de Santiago de Guayaquil (Ecuador)

[irene.trelles@cu.ucsg.edu.ec](mailto:irene.trelles@cu.ucsg.edu.ec)

## **MIRIAM RODRÍGUEZ BETANCOURT**

Universidad de La Habana (Cuba)

[justam@enet.cu](mailto:justam@enet.cu)

## **ANA TERESA BADÍA VALDÉS**

Universidad de La Habana (Cuba)

[ateresacu@yahoo.es](mailto:ateresacu@yahoo.es)

### **RESUMEN**

Se valoran tendencias sobre la conceptualización de la comunicación de la ciencia, la tecnología y la innovación, en un estudio teórico, cualitativo, desde la interpretación hermenéutica. Se muestran enfoques y conceptos que intentan explicar este tipo de comunicación y reflejan la fragmentación del pensamiento acerca de este fenómeno, lo cual se explica por su carácter emergente. El objetivo general es analizar algunos enfoques y conceptos sobre la comunicación de la ciencia, la tecnología y la innovación, dado el carácter polisémico de este concepto emergente. Se comparte con Jorge Huergo (2001) el abordaje de la comunicación, no entendida solamente en su dimensión mediática, sino como proceso que se hace comprensible en relación con la cultura, y por ello, se trata de la construcción de significados

compartidos. Se identifica en el trabajo la diferencia en cuanto a la comunicación de la ciencia de la tecnología y la innovación, asumida como divulgación de la ciencia; la concepción que la relaciona con los niveles de comunicación propuestos por Martín Serrano, y permite ubicarla en el nivel mediático, y finalmente, se analiza la concepción dialógica y democrática concebida en el modelo de Freyre (1973), al cual se adhieren las autoras.

#### **PALABRAS CLAVE**

Comunicación de la ciencia; lenguaje científico.

*Science communication; scientific language.*

## **Introducción**

El estudio de la comunicación de la ciencia, la tecnología y la innovación presenta, en el mundo contemporáneo, signado por cambios tecnológicos que se producen a ritmos exponenciales, una vital importancia.

Resultaría obvio aseverar que aquellos capaces de aplicar nuevas tecnologías resultantes del desarrollo del conocimiento científico, se encontrarán en posiciones mucho más ventajosas para adaptarse a los cambios continuos en una sociedad de la incertidumbre, y, aunque en modo alguno puede afirmarse que ello depende exclusivamente de la comunicación, pues se trata de un fenómeno multidimensional, con fuertes componentes de mediaciones desde lo económico, lo político, lo social, el desarrollo del conocimiento y la cultura científica, por solo mencionar algunos, corresponde un papel protagónico a la comunicación para lograr la socialización del conocimiento a nivel social.

En este trabajo se presenta una sistematización de tendencias conceptuales en torno al tema, como resultado de un estudio teórico, de corte cualitativo, desde la interpretación hermenéutica.

Los resultados dan cuenta de la diversidad de enfoques y conceptos que intentan explicar este tipo particular de comunicación y reflejan la fragmentación del pensamiento acerca de este fenómeno, lo cual se explica por el carácter emergente de esta disciplina.

## **Comunicación de la ciencia, la tecnología y la innovación Aproximaciones a la polisemia del concepto y enfoque desde la teoría social de comunicación de Martín Serrano**

Entre los primeros hallazgos que se producen al iniciar el estudio de este tema se encuentra el descubrimiento de su carácter polisémico, condicionado quizás por su carácter de disciplina emergente.

En tal sentido, el investigador y profesor mexicano JUAN TONDA (2005, p. 41), afirma que la divulgación de la ciencia, la comunicación pública de la ciencia, y la difusión de la ciencia, “pueden entenderse en sentido amplio como sinónimos de la misma actividad”; sin embargo, hay autores cuyos criterios al respecto difieren de los que sostiene el autor mexicano.

Dada la diversidad de puntos de vista sobre los matices que diferencian estos conceptos, se hace necesario profundizar en ello.

JORGE HUERGO (2001, p. 10) parte del concepto de comunicación, “no entendida sólo como un fenómeno vinculado con los medios (ya sean masivos o grupales de comunicación), sino considerada fundamentalmente como un proceso que se hace comprensible en relación con la cultura”. El autor establece que la comunicación tiene una dimensión transitiva, y es esa dimensión la que la acerca al concepto de divulgación, al afirmar: en este sentido transitivo, comunicación está íntimamente emparentada con “divulgación”, ya que se supone que “comunicar” es transmitir al vulgo (*divulgare*), generosamente, algo que un actor o un sector social especializado posee y ha construido.

Para Huergo, por tanto, la divulgación de la ciencia, la tecnología y la innovación presenta un carácter trasmisivo, unidireccional, no dialógico, y es justamente ello lo que la diferencia de la comunicación.

Al profundizar en las implicaciones de este enfoque, resulta pertinente acudir a la propuesta de Martín Serrano en su modelo dialéctico de comunicación que incorpora la consideración de los niveles involucrados en estos procesos, niveles definidos por este autor como: interpersonal, grupal, organizacional y mediático. En esos niveles, el mediático, se caracteriza por un mayor nivel de unidireccionalidad, pero tal criterio debe ser manejado con un enfoque en absoluto esquemático, ya que los restantes niveles *per se* no garantizan el enfoque dialógico, aunque éste sí tiene muchas más posibilidades de expresarse.

Puesto de ese modo, podría pensarse entonces en que uno de los niveles más frecuentemente utilizados para la divulgación de la ciencia, la tecnología y la innovación, es el nivel mediático. Y, al emplearse otros niveles como el interpersonal o el organizacional, el posible enfoque bidireccional marcaría una diferencia que la acerca mucho más al concepto de comunicación de la ciencia y al de construcción de significados en los diversos campos científicos o tecnológicos.

La comunicación de la ciencia, la tecnología y la innovación debe ser concebida como un proceso dialógico, de construcción compartida. Desde la comunicación de la ciencia se construyen valores de cultura científica que se comparten en contextos concretos.

Por ello, en los procesos de comunicación científica en los que participa la academia, también debe participar la sociedad porque, de lo contrario, se asume, de hecho, un concepto elitista, íntimamente relacionado con el de capitalismo cognoscitivo que se traslada paternalmente a un público ignorante y desconocedor, como apuntaba Huerta. Y que puede de igual modo acercarnos al concepto de capitalismo cognitivo, en lugar de concebir el conocimiento como bien público por excelencia.

En tal sentido HUERGO (2001) relaciona el enfoque transitivo con el poder, en la medida en que uno es el que habla y otro el que escucha; un emisor ilustrado transmite y el vulgo destinatario de esa transmisión, recibe. A lo que une la presencia implícita del poder, desde el punto de vista del actor protagónico en el proceso de conocimiento (en este caso, científico, tecnológico de innovación) en tanto el que recibe es ignorado en todo proceso de conocimiento significativo. El autor relaciona estos modos de asumir la comunicación y la divulgación con lo que Paulo Freire denominaba alienación de la ignorancia (FREIRE, 1969; 1973). Se establecen a nivel social dos roles, por una parte, la mayoría, ignorante, que no sabe y que, además, no puede hacer un proceso de conocimiento, y por otra parte, una minoría poseedora del saber, capaz de trasmitirlo (HUERGO, 2001).

Para el citado autor, el concepto de divulgación científica presenta tres perspectivas que se producen en diversos momentos o etapas que caracterizan su desarrollo: una primera perspectiva, que ubica en el iluminismo, tomando como fuente a Habermas, caracterizado por una racionalidad apoyada en tres esferas autónomas: la ciencia, el arte y la moral, organizadas en torno a aspectos específicos de su validez. Pero, por otro lado, y para que la racionalidad en ascenso pudiera legitimarse, debió realizarse una profunda distinción entre la cultura de los expertos y la cultura del público. Articulada con el desarrollo del capitalismo, esta perspectiva deseaba emplear esa acumulación de cultura especializada en la organización racional de la vida social cotidiana, para promover el control y dominio de las fuerzas naturales, un modo de comprensión del mundo y del individuo y la felicidad de todos los hombres (HUERGO, 2001, p. 8).

Una segunda perspectiva, a la que identifica con la “tradicción escolarizadora”, en la cual “existe una institución (articulada con el capitalismo y los rastros iluministas) destinada a transmitir los productos de la ciencia moderna y a llevar moralización, orden y control, en una palabra: ‘normalidad’, a las prácticas cotidianas desordenadas” (HUERGO, 2001, p.12).

Se coincide en el enfoque crítico de esta segunda perspectiva, que revela la copia e importación de significados y valores que signó a esa tradición escolarizada en América Latina, y que supuso volver la espalda a las culturas ancestrales y el conocimiento propio y auténtico, en lugar de inspirarse en el pensamiento martiano de defensa de la identidad, al decir que se debía injertar el mundo en nuestras repúblicas, pero que el tronco había de ser de nuestras repúblicas.

Otra crítica a formular a esta perspectiva, es que devela su naturaleza mecánica y formal en su determinación de lo que es el conocimiento científico y lo que se considera como tal en la tradición escolarizadora, al considerar un solo tipo de conocimiento, el denominado científico, que es el único llamado a transmitirse en la escuela. En tal sentido, sería necesario esclarecer entonces, que los contenidos no son sólo “científicos”, sino que representan nudos significativos del saber social y que, además, en cuanto “científicos” no son más que arbitrarios, es decir, transposiciones arbitrarias de las producciones científicas, cuyos filtros suelen ser: el programa, el manual y la propia construcción que de ellos hace el docente (HUERGO, 2001).

A ello se suma, el carácter bancario de esa postura, al decir de FREYRE (1969), quien criticó la educación y su divorcio con los valores de la cultura popular y propia, educación orientada, dice HUERGO (2001), a “un ideal de ciudadano idóneo compatible con el ideal burgués capitalista (TABORDA, 1951), en perjuicio del sujeto político autónomo” (p. 15).

La tercera y cuarta perspectivas, que identifica y analiza críticamente el mencionado autor, se relacionan con el difusionismo desarrollista, orientado a “difundir generosamente la racionalidad y la cultura modernizadoras de las naciones y los sectores nacionales desarrollados, a las naciones y sectores nacionales subdesarrollados, persistentemente tradicionales” (p. 16), con la finalidad de desarrollar al máximo posible el consumo de productos y tecnologías.

La mirada crítica del autor revela puntos comunes en estos conceptos de divulgación científica, entre los cuales destaca la relación que presentan entre divulgación y extensión, y citando a Paulo Freire, afirma: “el término extensión se encuentra en relación significativa con transmisión, entrega, donación, mesianismo, mecanicismo, invasión cultural, manipulación (FREIRE, 1974); es decir, con el significado transitivo de comunicación” (HUERGO, 2001, p. 16).

Aboga en cambio por un concepto opuesto a la divulgación y tendiente a la comunicación en el sentido de construcción compartida del conocimiento científico y tecnológico, considerando la ciencia como proceso social que atañe a todos los sectores y que se establece en relación con diferentes culturas (DELGADO & QUEVEDO, s. f.), sino también la comprensión del otro, no ya como masa indiferenciada de no-especialistas o de simples observadores pasivos, sino más bien sujeto social y de conocimiento; lo que implica considerar al otro como sujeto de comunicación, y no como objeto o destinatario; un sujeto que debe ser involucrado en un proceso colectivo de conocimiento, razonamiento, pensamiento y crítica acerca de cuestiones científicas y tecnológicas (FEHÉR, 1990). Lo que implica, por otro lado, trabajar con el otro, y no para el otro, produciendo una ruptura en las relaciones basadas en la autoridad cognitiva, en la dominación del especialista o en la manipulación técnico-científica.

Como propuesta para transformar tal orden de cosas, el citado especialista propone el modelo de Martín Barbero de pensar y comprender la comunicación desde las mediaciones, y lo toma como base para su concepto de *popularización de la ciencia*, en lugar del de divulgación científica.

La popularización de la ciencia es un proceso dialógico, que, a juicio de HUERGO (2001) consiste en un persistente involucramiento de los sujetos en los procesos colectivos de conocimiento, razonamiento, pensamiento y crítica acerca de cuestiones científicas y tecnológicas; y además de ellos, debe conducir a una permanente democratización de las producciones científico-tecnológicas, esto es: trabajar en favor del acceso, la apropiación y el uso de las producciones alcanzadas por una sociedad, que dejan de ser individuales o particulares, para convertirse en colectivas. La comunicación dialógica no sólo

gira en torno a formas, a ocasionales encuentros espontáneos, sino que se refiere a contenidos, a producciones sociales, sobre las cuales se basan los procesos de apropiación y de producción de significados.

Por tanto, de la misma manera que para algunos el proceso de investigación científica incluye el de dar a conocer los resultados obtenidos, para este autor, el proceso de popularización de la ciencia involucra no sólo el proceso de comunicarla, sino también el de democratización, apropiación y uso de la producción científica. Con lo cual ciertamente asume un enfoque democrático y avanzado sobre la producción científica y su popularización, a nivel macro social, que compartimos.

Su materialización pasa por la adopción de políticas públicas de ciencia, tecnología y educación tendientes a la democratización del conocimiento que a su vez sirven de marco y median en acciones comunicativas que deben realizarse a otros niveles, como el interpersonal, el grupal, el organizacional o institucional, el interinstitucional y el mediático. Acciones de construcción de significados y sentidos en contextos sociales, históricos, culturales que los median y a su vez son mediados por ellos.

Estrechamente relacionado con la interpretación de la popularización de la ciencia, propuesto por HUERGO (2001) y analizado en extenso, pueden encontrarse otros como apropiación del conocimiento científico, o socialización de éste.

En tal sentido, define MARTÍN AGUDELO (2012, p. 57): “la búsqueda de la sociedad del conocimiento ha dado lugar a que las comunidades científicas proyecten estrategias encaminadas a socializar o ‘popularizar’ los conocimientos obtenidos de sus investigaciones para que los diferentes actores de la sociedad logren comprenderlos y asimilarlos”.

A esto se le ha denominado apropiación social del conocimiento. La apropiación social del conocimiento, a juicio de MARTÍN AGUDELO, consiste en un proceso que, por un lado, “dispone los conocimientos científicos y tecnológicos en un escenario y lenguaje comunes para la sociedad; y por otro, que el ser

humano hizo suyos tales conocimientos, como elementos útiles y necesarios para su beneficio y provecho” (2012, p. 57).

Este concepto presenta numerosos puntos de contacto con el concepto de popularización de la ciencia analizado anteriormente, pues al igual que aquel apuesta por una democratización del conocimiento a nivel macro. Dicho con palabras del autor, significa, entonces: la democratización del acceso y uso del conocimiento científico y tecnológico, como estrategia para su adecuada transmisión y aprovechamiento entre los distintos actores sociales, que derivará en el mejoramiento de la calidad de vida de las comunidades y sus integrantes. Significa un reto epistémico en cuanto al modo dominante de producción de conocimientos en las universidades, porque estos no deben quedar reducidos a las elites intelectuales (p. 57).

Entre las áreas, profesiones y disciplinas que de una u otra forma contribuyen al mejor aprovechamiento de los conocimientos científicos en beneficio de la comunidad, cita el autor la comunicación y el periodismo. Destaca MARTÍN AGUDELO (2012) que “las actividades de divulgación y apropiación social del conocimiento en los países de Latinoamérica han recaído esencialmente en los medios de comunicación y, por ende, en los comunicadores y periodistas” (p. 58).

Pero el ejercicio profesional en este campo en opinión del colombiano no se presenta libre de contradicciones y problemáticas, pues señala críticamente, por un lado, las limitadas oportunidades de acceso a los medios de comunicación y la extrema concentración de estos en las grandes ciudades, y por otro, el hecho de que el peso fundamental de estas actividades, que constituyen “tareas estratégicas del desarrollo científico y tecnológico de la región recae solo en los mediadores y no en los generadores de conocimiento” (AGUDELO, 2012, p. 58).

Citando a Lozano, apunta: El periodismo científico, que ha servido como fuente de aprendizaje, busca hacer comprensibles los conocimientos derivados de estas investigaciones y constituye una verdadera herramienta de alfabetización científica, porque está dirigido a públicos amplios con diferentes

niveles educativos, económicos y sociales. Sin embargo, faltan, tanto en cantidad como en calidad, profesionales dedicados a la comunicación, divulgación o popularización de la ciencia, que sepan interpretar y reinventar las representaciones que de ella se crean en el imaginario de los ciudadanos, para que puedan apropiárselas (LOZANO, 2005, citado en MARTÍN AGUDELO, 2015, p. 58).

## **La comunicación de la ciencia, la tecnología y la innovación y el periodismo científico**

Las autoras de este trabajo son del criterio de la necesidad de promover la “gestión de comunicación de la ciencia y tecnología, entendida como socialización del conocimiento científico, y de manera paralela continuar el debate académico en cuanto a aproximaciones y conceptos” (TRELLES & RODRÍGUEZ, 2013).

El periodismo científico se ha entendido como aquella especialización informativa que consiste en divulgar la ciencia a través de los medios de comunicación. Una definición más completa la encontramos en la concepción de la especialista venezolana ARGELIA FERRER (2003), quien lo describe como el que “selecciona, reorienta, adapta, refunde un conocimiento científico, producido en el contexto particular de ciertas comunidades científicas, para que ese conocimiento transformado pueda ser apropiado dentro de un contexto distinto y con propósitos diferentes por una determinada comunidad cultural” (p. 14).

Sus objetivos, por tanto, se fundamentan en hacer accesible el conocimiento y poner al alcance de la mayoría el patrimonio científico de una minoría.

El periodismo científico, como lo conocemos hoy, comenzó en el siglo XX, específicamente de los años 20, cuando Waldemar Kaempffert inició una sección de crónicas científicas en el *New York Times*. Sin embargo, la temática de salud, como área de comunicación especializada, ya aparece en los siglos XVIII y XIX.

Definido, pues, como la especialización informativa encargada de divulgar la ciencia a través de los medios de comunicación, hay que inscribirlo, por consiguiente, en lo que se denomina Periodismo Especializado, que, según la profesora española MONTSERRAT QUESADA (2004), es “el que resulta de la aplicación minuciosa de la metodología periodística de investigación a los múltiples ámbitos temáticos que conforman la realidad social, condicionada siempre por el medio de comunicación que se utilice como canal” (p. 133).

De las diferencias entre Periodismo Especializado y el Periodismo Generalista o Informativo se infieren los rasgos que sustentan la inclusión del Periodismo Científico como especializado ya que, contrario al generalista, trasciende la urgencia informativa, le interesa explicar las causas de los hechos, de ahí que requiera fuentes especializadas y variadas, periodo más extenso de elaboración, mayor espacio y/o tiempo y mejor preparación previa. Existen diferencias además con relación a los métodos de investigación empleados, que en el Especializado son de mayor rigor, con más amplia documentación, pero no son distintos en lo tocante a la ética, la objetividad, y la propiedad en el lenguaje.

Nada impide al profesional que ejerce este tipo de periodismo comentar, analizar y reflexionar sobre las materias que aborda, sea en una información, un artículo o cualquier otro género periodístico, y, por supuesto, contextualizarlas, para lo cual requiere de conocimientos sólidos acerca del tópico científico en cuestión y de formación periodística especializada.

Este profesional necesita, en consecuencia, preparación relacionada con la comunicación, la ciencia, la tecnología, el periodismo especializado, la redacción periodística, entre otras materias que le posibiliten comprender en teoría y práctica las relaciones comunes entre los discursos mencionados, y sus relaciones distintivas.

Como coinciden varios expertos, el periodismo científico es una actividad esencialmente educativa porque se propone, entre sus objetivos esenciales, popularizar el conocimiento científico y tecnológico para vencer el oscurantismo que atenta contra el desarrollo humano, a la vez que instar a las

instituciones y decisores pertinentes a que se introduzcan los resultados científicos en la sociedad. Lograr estos objetivos sería prácticamente imposible sin el concurso de los medios de comunicación.

Ciencia que no se comunica, no existe, expresó reiteradamente con su habitual lucidez el famoso y reconocido divulgador norteamericano de la ciencia, CARL SAGAN. Ninguna máxima más certera para resaltar la importancia del Periodismo Científico, la especialización que se sustenta en los medios masivos de comunicación, al decir de ZAMARRÓN (2006, p. 136): “tan determinantes en la vida pública actual, para conformar opinión en las personas y en captar la atención de niños y jóvenes, que se han convertido en ‘la otra escuela’ más ágil, inmediata, convincente, espectacular y de por vida”.

Al comunicar masiva y simultáneamente los procesos, hallazgos y progresos científicos y tecnológicos, además de familiarizar al público con terminologías propias de este ámbito, se facilita la adopción de decisiones y se contribuye a la formación de juicios de valor, los que deben traducirse en actitudes, como exponen los especialistas en comunicación.

Otro de los roles u objetivos fundamentales que debe asumir el periodismo científico es el de relacionar los avances de la ciencia con la vida cotidiana de las personas, de tal modo que éstas la perciban como espacio cercano, comprensible y, sobre todo, imprescindible para su propio desarrollo. Para ello, el periodista debe referirse a cómo se genera la investigación y el consiguiente descubrimiento, además de concientizar al receptor acerca del contexto social, económico y político de la ciencia y de su impacto en la sociedad.

Dado el mayor conocimiento de esta temática por el receptor contemporáneo, al menos en lo que respecta a sus logros y aplicación de ellos en la vida social, la multiplicidad de canales divulgativos y los espectaculares avances científicos que se registran actualmente, el periodismo científico adquiere una connotación especial. Al respecto, MANUEL CALVO HERNANDO (2005), el prestigioso experto español en la temática de la Sociedad de la Información, ha señalado que “la información científica se convierte en una sugestiva y a veces divertida o escalofriante caja de sorpresas” (p. 27).

Estas circunstancias imponen nuevos retos, especialmente en cuanto a la concepción y elaboración de los productos comunicativos y, por supuesto, respecto a las funciones y objetivos de los profesionales. Se hace necesario, en tal contexto, repensar tanto los modos y estrategias para diseñar mensajes atractivos como los roles que los periodistas deben asumir en este trabajo.

A pesar del conocimiento sobre determinados aspectos de la ciencia que, en efecto, han adquirido amplios públicos, y, mayormente, audiencias selectivas, aún existe una gran ignorancia en el mundo sobre los problemas en esta esfera y sus impactos en la vida cotidiana, lo que se agrava en muchas ocasiones por el grado de especialización de cada disciplina.

En este sentido, muchos expertos han mostrado su preocupación sobre el peligro de la tendencia a la superespecialización defendida por un sector del campo, y se decantan por un equilibrio a partir del principio básico que debe ser inherente a la Comunicación de la Ciencia: acercar los saberes de la minoría al conocimiento de los sectores más amplios.

## **Algunos desafíos de la comunicación del periodismo científico**

El Periodismo Científico afronta retos y obstáculos de diverso carácter, uno de los más difíciles se encuentra en aquellos países —la mayoría— donde son débiles los sistemas de investigación y desarrollo, y escasa la valoración de la ciencia como elemento constitutivo de la cultura por lo que tampoco se presta la debida atención a la formación de los profesionales en esta especialización.

Paradójicamente, es en dichas regiones donde existe imperiosa necesidad de acceder a los conocimientos científicos que puedan ayudar a sus pobladores a afrontar problemas cotidianos, en torno a los cuales el PEC, por dirigirse a públicos heterogéneos y emplear diversos canales y recursos expresivos, puede prestar un servicio eficaz, de conjunto con otros instrumentos de la comunicación.

Un problema clave sigue siendo el lenguaje con el que se pretende llegar al público, diverso o conocedor, sea por un canal u otro, pues los códigos científicos tienen, en muchos casos, particularidades que deben ser respetadas por los comunicadores para no caer en tergiversaciones. No traicionar ni simplificar el lenguaje propio de estas disciplinas y, al mismo tiempo, hacer comprensible los mensajes al receptor con los términos que faciliten la comunicación es uno de los retos más difíciles, y al mismo tiempo, más estimulantes para los periodistas especializados.

Para comunicar ciencia a públicos heterogéneos mediante los medios masivos, el periodista científico dispone de un instrumento expresivo *sui generis*: el lenguaje periodístico, caracterizado por la claridad, la síntesis y la concisión entre sus atributos esenciales.

Es un lenguaje que requiere el máximo de recursos y formatos expresivos para contar e interpretar lo que pasa día a día en el mundo, especialmente en el ámbito de la ciencia que, a su vez, se basa en su propio código lingüístico. El dominio de los géneros periodísticos deviene aspecto de especial relevancia lo mismo en publicaciones impresas que en las emisiones audiovisuales y, por supuesto, en la red.

El periodista escribe inevitablemente en géneros que son categorías de obras definidas por reglas comunes y características semejantes. Estas obras cumplen determinadas funciones, como las de informar, trasladar opiniones, contribuir a la concientización, destacar hechos y figuras, entre las más importantes.

Aunque la hibridación de géneros, su entrecruzamiento y la convergencia de medios constituyen hoy realidades innegables, en la mezcla no pierden sus funciones esenciales, de modo tal que al examinarlas podemos definir a qué género periodístico concreto responde determinada composición en cualquier medio de comunicación.

A los géneros periodísticos más antiguos —la nota informativa o información, la entrevista, el reportaje, el artículo, la columna, la crónica— ha ido

añadiéndose un considerable repertorio que incluye la documentación, el dossier, el análisis, el informe, muy empleados en periódicos, revistas y publicaciones similares.

Conociendo bien las técnicas de redacción (exposición, diálogo, narración y descripción), el periodista científico elegirá entre el mencionado arsenal genérico cuál es la estructura más propicia para revelar descubrimientos, aclaraciones y opiniones sobre los asuntos que trate.

Para dar información puntual de descubrimientos, se emplea la información, por su concisión, atribución de fuentes, y la inmediatez publicística que se le otorga en dichas publicaciones, pero si el periodista se propone, por ejemplo, explicar un problema o asunto de alta especialización, tal vez resulte más adecuado entrevistar al científico responsable, instándole, además, a que lo haga del modo más claro que le sea posible.

Parecería ideal abordar un tema científico o tecnológico controvertible mediante el reportaje, en el que las opiniones divergentes pudieran tener cabida, o elaborar un dossier para mayor completitud. Desde luego, para aplicar soluciones como estas se requiere que el periodista científico tenga un conocimiento cabal de las características y funciones de cada género y también del perfil editorial del medio en el que insertará sus exposiciones, entre otras variables de la complejidad mediática.

En el caso de los medios audiovisuales, el empleo de los géneros alcanza una dimensión extraordinaria habida cuenta del impacto de la imagen y del desarrollo tecnológico que en este terreno se ha alcanzado. Auxiliada con los recursos propios del lenguaje audiovisual, la información puede contribuir al referir un descubrimiento, a que el público perciba la utilidad inmediata de este hallazgo para beneficio propio o de la sociedad, pues se precisa cómo el hecho tiene repercusión concreta en la vida social.

De igual forma, el reportaje en televisión, por su carácter impresionista, poniendo en pantalla a los protagonistas o testimoniados, y a los escenarios, da una dimensión más objetiva al público del asunto tratado.

La entrevista, en conversación con autores, testigos o expertos relacionados con cualquier aspecto de la ciencia, logra establecer un diálogo muy especial entre las fuentes y el receptor, y contribuye, además, a socializar la imagen de los científicos.

Especial significación cobra el comentario casi siempre expuesto por un experto, para llamar la atención del televidente sobre un hecho específico, el que puede también ser graficado con el fin de reforzar el efecto persuasivo y atraer y mantener la atención del receptor.

En cuanto a la elaboración y redacción de artículos de carácter científico, en cualquiera de sus modalidades y para cualquier medio que tenga como destinatarios a audiencias generales, requiere particular atención debido al carácter mismo de estas obras en lo que se refiere al tono y lenguaje de su discurso.

Como dicen CISNEROS y OLAVE (2012), “el artículo científico se adecúa generalmente a la tipología de texto expositivo-argumentativo” (p. 18), aunque la redacción de carácter periodístico, por la heterogeneidad de sus públicos y canales, posibilita que los trabajos puedan expresarse en una amplia variedad de géneros, los más tradicionales, como la nota informativa, la entrevista, el reportaje, la crónica, el ensayo, y otros que han ido surgiendo, ya con patente de permanencia, digamos el dossier, el informe, el análisis, entre otros.

Si bien cada medio tiene su lenguaje y características definidas, y en dependencia de ello los especialistas componen los trabajos periodísticos de carácter científico –al igual que los de cualquier otro tema–, lo que les es común a todos para asegurar el éxito comunicativo es la escritura de calidad, en tanto todo producto mediático es, primero, producto escrito o, por lo menos, ideado en palabras. Esto vale igualmente para la red, que ha abierto ilimitadas posibilidades para la elaboración y difusión de mensajes de todo tipo y naturaleza mediante la interactividad, los enlaces, la integración de gráficos, imágenes y sonido, entre otras técnicas y recursos que generan, a su vez, nuevos géneros.

Y, por supuesto, vincularse directamente con las fuentes primarias, es decir, con los científicos que, por su conocimiento y protagonismo, son los llamados a explicar los elementos de un descubrimiento o aclarar procesos que, sin su experticia, sería imposible entender y, mucho menos, describir.

El éxito de la comunicación de la ciencia con el empleo de los géneros periodísticos reside en el convencimiento por parte del comunicador, realizador, del periodista, de que “comunicar ciencia, y comunicarla bien, constituye uno de los grandes retos que afronta el comunicador de nuestros días. Y que ese reto sólo puede ser asumido y vencido con preparación, rigor y respeto de los lenguajes propios tanto de la ciencia como de la comunicación” (RODRÍGUEZ, 2006, p. 220).

## **A manera de conclusiones**

La comunicación de la ciencia, la tecnología y la innovación ha adquirido una relevancia que aumenta de manera gradual en el mundo contemporáneo, aunque debe destacarse que no ha logrado aún conceptos consensuados en cuanto a su denominación, naturales, objeto de estudio.

Entre las tendencias que intentan explicar este campo se advierten las relacionadas con modelos de comunicación que de manera implícita revelan las concepciones en relación con el protagonismo de los sujetos y la democratización de los procesos tanto de producción del conocimiento científico, tecnológico y de innovación, como de su comunicación, que ha de articular a la esfera mediática en permanente concepción dialógica.

Las autoras del presente trabajo apuestan por las concepciones que defienden la necesidad de democratización de esos procesos, pues únicamente así podrían lograrse las ambiciosas metas de enriquecimiento humano y social que el siglo XXI demanda.

## Bibliografía

- CALVO, M. (2005). Desafíos del Siglo XXI para la divulgación de la Ciencia: globalidad, complejidad y expansión incontrolada del saber. En Antonio Marín, Irene Trelles, y Guadalupe Zamarrón, G. (compiladores). *Universidad y Comunicación Social de la Ciencia* (pp. 21-44). Granada, España: Editorial Universidad de Granada.
- CISNEROS, M.; OLAVE, G. (2012). *Redacción y publicación de artículos científicos. Enfoque discursivo*. Bogotá, Colombia: Ecoe Ediciones.
- CRUZ, J. (2016). Los elementos del Periodismo. [http://politica.elpais.com/politica/2016/09/12/actualidad/1473706925\\_208759.html](http://politica.elpais.com/politica/2016/09/12/actualidad/1473706925_208759.html)
- FERRER, A. (2003). *Periodismo Científico y Desarrollo. Una mirada desde América Latina*. Mérida, Venezuela: Ediciones del Rectorado.
- FREIRE, P. (1973). *El mensaje de Paulo Freire. Textos seleccionados por el INODEP*. Fondo de Cultura Popular. Madrid: Ed. Marsiega.
- HABERMAS, H. (2012). *Mundo de la vida, política y religión*. Trotta: Madrid.
- HUERGO, J. (2001). La popularización, mediación y negociación de significado. La popularización de la Ciencia y la Tecnología. En: Seminario Latinoamericano Estrategias para la Formación de Popularizadores en Ciencia y Tecnología. Red-POP - Cono Sur. La Plata, 14 al 17 de mayo de 2001 Red-POP. <http://admin.redpop.org/redpopAsp/paginas/>
- MARÍN AGUDELO, S A; (2012). Apropiación social del conocimiento: Una nueva dimensión de los archivos. *Revista Interamericana de Bibliotecología*, 350, 55-62. <http://www.redalyc.org/articulo.oa?id=179024991005>
- MARTÍN SERRANO, M. (2004) Presentación de la Teoría Social de la Comunicación. En: *La producción social de comunicación*. Alianza: Madrid.
- QUESADA, M. (2004). Periodismo de Investigación: una metodología para el Periodismo especializado. En *Periodismo Especializado*. Fernández del Moral (coord.) Barcelona: Editorial Ariel.
- RAYA, M.; ZULUETA, M. (2011). *Textos científico-técnicos. ¿Cómo crearlos?* La Habana: Instituto Cubano del Libro.
- RODRÍGUEZ, M. (2005). Expresiones periodísticas en la divulgación de la Ciencia y la Técnica. En: *La comunicación de la Ciencia y la Tecnología. Una visión universitaria*. La Habana: Dirección de Extensión Universitaria. pp.116-128.

- RODRÍGUEZ, M. (2006). Comunicación de la Ciencia. Algunas ideas básicas. En UNIVERSIDAD, COMUNICACIÓN Y CIENCIA: CONTRASTES. Marín, A. Pacho, C., Trelles I. y Zamarrón, G. (Coordinadores). Universidad Autónoma de Baja California. Mexicali. Departamento de Editorial Universitaria. pp. 213-220.
- TONDA, J. (2005). ¿Qué es la divulgación de la Ciencia? En *La Comunicación de la Ciencia y la tecnología. Una visión universitaria*. La Habana, Cuba: Ministerio de Educación Superior. pp. 40-48.
- TRELLES, I.; PACHO. C. (2006). El estudio de la imagen universitaria en medios. En *Universidad, comunicación y ciencia: contrastes*. Marín, A. Pacho, C. et al. Universidad Autónoma de Baja California. Mexicali: Departamento de Editorial Universitaria. pp. 95-108.
- TRELLES, I.; RODRÍGUEZ, M. (2013). Comunicación de la Ciencia y la Tecnología y cultura científica para la prevención de riesgos. Experiencia cubana, en Miguel Hernández Communication Journal, 10, 213-231. Universidad Miguel Hernández, UMH (Elche-Alicante). <http://mhcj.es./index.php?journal=mhcj&page=article&op=view&path>
- TRÉSPIDI, M. (2005). Cultura Ciencia y Universidad en Latinoamérica. Reflexiones Generales y Estudio de un caso. En *Universidad y Comunicación Social de la Ciencia*. Marín, A, Trelles I. y Zamarrón G (coordinadores). Granada: Editorial Universidad de Granada. pp. 45-71.
- VALSAGNA, A. (2006). La implementación de un programa de imagen y comunicación institucional en una universidad pública. El caso de la Universidad Nacional del Litoral, Argentina. En *Universidad, comunicación y ciencia: contrastes*. Marín A., Pacho C, et al, (coordinadores). Mexicali, Universidad Autónoma de Baja California. pp. 75-94.
- ZAMARRÓN. G. (2006). La divulgación universitaria de la ciencia: entre el deber y el aplauso. En *Universidad, comunicación y ciencia: contrastes*. Marín, A. Pacho. C et al. Coordinadores. Universidad Autónoma de Baja California. Mexicali: Departamento de Editorial Universitaria. pp. 129-145.
- ZAMARRÓN, G. (2006). De cultura científica y anexas. En *Universidad, comunicación y ciencia: contrastes*. Marín, A., Pacho C. et al. (Coordinadores). Universidad Autónoma de Baja California. Mexicali. Departamento de Editorial Universitaria. pp. 129-145.

# De la gestión de archivos sonoros: teorías y conceptos

## **ANA TERESA BADÍA VALDÉS**

Universidad de La Habana (Cuba)

[ateresacu@yahoo.es](mailto:ateresacu@yahoo.es)

## **IRENE TRELLES RODRÍGUEZ**

Universidad Católica de Santiago de Guayaquil (Ecuador)

[irene.trelles@cu.ucsg.edu.ec](mailto:irene.trelles@cu.ucsg.edu.ec)

### **RESUMEN**

La mirada a los archivos sonoros debe realizarse desde la perspectiva teórica de la Ciencia de la Información- y de la Ciencia de la Comunicación, entre otras ramas de estudio. En la actualidad, gran parte de lo sonoro no se recopila de forma sistemática. Y es contradictorio que en esta era de tecnologías, en la cual las producciones digitales se pueden escuchar, almacenar, intercambiar y producir en plataformas novedosas, no existan muchas garantías ni certezas de que estén siendo archivadas y preservadas adecuadamente a nivel mundial. El desarrollo teórico y práctico del tema es novedoso pero muy poco extendido, como lo demuestra la escasa bibliografía en el ámbito internacional. Este texto tiene como objetivo sistematizar teorías y conceptos de la Ciencia de la Información y la Comunicación que tributan a la gestión de archivos sonoros, fuentes del patrimonio de nuestras naciones. Se utilizan tres métodos investigativos del nivel teórico: el método de análisis y síntesis, el cual se emplea para realizar el análisis de textos que abordan el tema; se hace uso del método histórico lógico para realizar

un recorrido por los antecedentes del tema; y se aplica el método sistémico para articular teorías y conceptos relacionados con el tema. Se emplea como técnica el análisis bibliográfico-documental y la triangulación de datos.

#### **PALABRAS CLAVE**

Archivos sonoros; preservación; cultura.

*Sound archives; preservation; culture.*

## **Introducción**

Si se sigue la alerta de RODRÍGUEZ (2020), para los finales de la próxima década se habría perdido cualquier contenido sonoro almacenado en formatos análogos. La digitalización de contenidos es inaplazable para la preservación a largo plazo. En la actualidad existe además una marcada relación entre los archivos sonoros y las estrategias de preservación, su aplicación y las políticas organizacionales (FLORES FERNÁNDEZ, GÁTICA MOLINA, GONZÁLEZ CORREA, & NÚÑEZ QUINTEROS, 2022), que no puede descuidarse.

Este texto resulta un acercamiento a la historia, conceptualización, teorías y evolución de los archivos sonoros, visión que ha de realizarse desde las visiones teóricas de la Ciencia de la Información y la Comunicación.

Si bien no existe una definición última de los archivos sonoros, se les define como parte de lo audiovisual, esto último referido:

a las imágenes en movimiento y/o a los sonidos grabados, registrados en película, cinta magnética, disco o cualquier otro medio actualmente conocido o por inventar (...). Archivo quiere decir una organización o unidad de una organización que se dedica al acopio, la gestión y la preservación de una colección de materiales audiovisuales y afines, facilita el acceso a la misma y su utilización. El término abarca organizaciones oficiales y no oficiales, comerciales y culturales que cumplen estas cuatro funciones (EDMONSON, 2004, p. 85).

A través de las diferentes épocas, los archivos audiovisuales han empleado una denominación concreta heredada del término francés *bibliothèque* y el español 'biblioteca'; de ahí vienen *cinemathèque* (cinemateca, sinematek, kinemathek) para los archivos cinematográficos, *phonothèque* o discoteca para los archivos sonoros y *médiathèque* para los archivos de medios audiovisuales.

Para EDMONSON (op. cit.) el adjetivo audiovisual se encuentra "dirigido a los sentidos de la vista y el oído"; cada vez se emplea más como vocablo único utilizado para incluir por igual a las imágenes en movimiento y los sonidos grabados de todo tipo. Con cierta diversidad, el término figura en los títulos de algunos archivos y asociaciones profesionales al tiempo que es el que ha sido adoptado por la Organización de Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) para agrupar los campos de los archivos cinematográficos, de televisión y de sonidos. Desde esas perspectivas, el archivo sonoro resulta una organización signada por el acopio, la gestión, la preservación y la accesibilidad (RODRÍGUEZ, 2016). Esta última característica es de imprescindible comprensión en el entorno digital caracterizado por la participación de cada vez más activos prosumidores en los diferentes procesos. Igualmente resultan pertinentes en tal sentido las aportaciones referidas a las especializaciones temáticas archivísticas. Como ocurre con el resto de las profesiones vinculadas con la recopilación de documentos:

la especialización, tanto en temas como en medios, es una característica que se encuentra tradicionalmente en los archivos audiovisuales y sin duda no se modificará, por motivos relacionados tanto con las predilecciones personales como con el apoyo a las necesidades y objetivos institucionales. Por consiguiente, hay archiveros cinematográficos, de sonido (EDMONSON, op. cit., p. 67).

Los archivos sonoros forman parte de la información en la era digital (BADÍA, 2013). En ella se ha fortalecido el énfasis de las organizaciones en la producción corporativa para tomar ventajas en los cambios sociales y tecnológicos (PONJUÁN, 2008). Si se quisiera ilustrar la renovada presencia de lo sonoro a través de Internet, valga citar la existencia de sitios como iTunes,

eMusic, Allmusic, Spotify, Deezer, Muze, Amazon, Rhapsody y Lala, en los cuales se pueden escuchar miles de documentos de los más diversos géneros. Por otra parte, en las redes sociales, también existen sitios como Soundhound, MySpace, Lastfm y Soundcloud, que posibilitan a las audiencias publicar, escuchar e interactuar. Asimismo, existe una diversidad de estaciones radiofónicas que constantemente publican podcasts y otras producciones sonoras.

La implantación de las técnicas digitales ha llevado aparejado un proceso de transformaciones tecnológicas, económicas y sociales. Las primeras han descollado en la digitalización de todos los procesos sonoros: la automatización de la producción y transmisión; el almacenamiento digital; y la compresión de lo que se escucha, entre otros. Asimismo, se ensanchan las posibilidades de interacción, de diálogo y participación de las audiencias. Los cambios de la era digital permiten que las audiencias puedan escuchar ahora lo que desee en diferido sin estar sometidas a la linealidad radiofónica. Ese tipo de recepción favorece la personalización, una de las tendencias comunicativas del siglo XXI. Otra de las características más difundidas hoy es el almacenamiento de gran cantidad de documentos sonoros en cada vez más pequeños espacios.

En la archivística se puede identificar la ocurrencia de dos estadios paradigmáticos (MENA, 2015). El primero de ellos, manifestado desde el surgimiento de esa disciplina a fines del siglo XIX y hasta las postrimerías del siglo XX, se ha denominado según el enfoque de varios autores, custodial, patrimonialista, histórico-tecnicista, clásico, industrial, positivista, y estatista. En el caso del segundo, se ubica a partir de las décadas de los 80-90 del siglo XX y ha recibido las denominaciones de postcustodial, social, postcustodial científico informacional, del acceso, postmoderno o nuevo paradigma archivístico.

En esa “perspectiva evolutiva”, MALHEIRO DA SILVA (s. f.) analiza que “no puede descuidarse la valorización del contexto de producción, el análisis de los flujos y usos de la información, el estudio del comportamiento informacional, con

énfasis en las cuestiones psicológicas y cognitivas” (p. 5), a lo que habría que añadir también las comunicativas. Esa visión está modelada por rasgos como la valorización de la información como fenómeno social, el moviismo informacional frente al llamado “inmoviismo” documental, el acceso a la información, la comprensión y el conocimiento de la información social.

Todas esas aportaciones advierten los influjos de los Estudios Culturales, con su entendimiento de la cultura –y la comunicación– como el conjunto socialmente aprendido, construido, compartido y perdurable de valores, hábitos, costumbres, normas, reglas, códigos, impresiones, opiniones, creencias, conocimientos, actitudes, documentos, instituciones, hechos, y manifestaciones sociales en general. En ello no puede obviarse el influjo de los contextos y la participación. Se coincide con LINARES (2004) cuando alerta que la crítica del sociocognitismo al enfoque cognitivo ortodoxo, introduce la necesidad de apostar por las determinaciones sociales y culturales en cualquier propuesta conceptual en el terreno informacional.

En ese devenir no puede obviarse que la aparición del concepto de efecto cognitivo marcó a mitad de los 70 un giro en lo comunicacional. Es cuando más comienzan a tenerse en cuenta no los efectos puntuales –derivados de la exposición a cada mensaje– sino “los efectos acumulativos, sedimentados en el tiempo” (WOLF, 1987, p. 158). Las nuevas visiones empezaron a encauzar sus rumbos en asuntos como el papel que ocupan las instituciones culturales como los medios de comunicación y las rutinas de sus profesionales en el desarrollo del sistema social.

Esas diferencias entre el viejo y el nuevo paradigma en Comunicación se sintetizaron en que ya no se priorizaría los estudios de casos individuales, sino la cobertura global de todo el sistema mediático-comunicacional. Se consolidaría el uso de metodologías más integradas y complejas, así como la reconstrucción del proceso mediante el cual el individuo modifica su propia representación de la realización social. Y las nuevas interpretaciones se tradujeron entonces en lo radiofónico-sonoro, en la aparición de reflexiones teóricas de corte hermenéutico-culturalista, cuyos senderos abarcaron la

consideración de la arqueología del receptor, la antropología del sonido, la etnomusicología, la semiótica radiofónica y la antropología de la recepción.

La totalidad de esas corrientes ofrecen miradas renovadas al considerar el mundo sensorial del sonido-con-sentido, el ambiente acústico de la cultura, la construcción del mensaje y la idea de la transmisión cultural-sonora, entre otros aspectos, que se preocupan por la construcción de significados en la comunicación, mediante la asociación texto, sonido, cultura y contexto.

Con relación a lo sonoro, MALDEN (2011) insiste en que no sólo se trata de crear documentos y contenido, sino que cada institución debe saber manejar lo que genera. Por otra parte, explica que se debe trabajar en una política de documentos esenciales (*core records*) para evaluarlos en términos de su importancia, el valor del creador y el valor para la organización. A esa política se une la variedad de servicios tradicionales que un archivo debe ofrecer, y lo otro esencial es el acceso, facilitado por la catalogación e indexación. Finalmente, hay que cumplir con la demanda del uso de los contenidos, lo cual supone que esta información va a estar accesible. Es importante que el archivo conozca cuáles son los usos de la información, cuáles son las necesidades de la comunidad, de los usuarios y que se esté consciente de que las necesidades pueden cambiar constantemente. Otra relación que debe ser gestionada, es la que debe existir con los creadores de contenido y los equipos de producción.

## Desarrollo

### El sonido

El sonido es “el resultado de percibir auditivamente variaciones oscilantes de algún cuerpo físico, normalmente a través del aire” (RODRÍGUEZ, 1998, p. 23). Esas variaciones pueden ser percibidas por el oído. De esas posturas teóricas se infiere que el sonido es un fenómeno físico, pero también fisiológico, psicológico y comunicacional. De ello se desprende que no toda vibración del aire puede ser considerada sonido, sino aquella capaz de estimular el sentido auditivo. El sonido, como objeto de estudio de la acústica, puede ser definido,

entonces, como las vibraciones producidas por objetos en movimiento que son transmitidas a través de ondas que se propagan, por distintos medios, principalmente el aire y que pueden ser percibidas por el oído. Cuando esas vibraciones son convertidas en señales mecánicas, eléctricas o digitales se pueden grabar y con ello fijar el sonido en un determinado soporte, dando lugar a la creación de un documento sonoro.

La expresividad de esos sonidos puede concretarse si sintetizamos la tipología de CEBRIÁN HERREROS (1993): narran el ambiente de una situación y enseñan la realidad, sirven de fondo para la información, pueden convertirse en información y contrapuntean palabras, música, efectos sonoros y silencio. Desde esa base, las principales funciones de lo sonoro –también en lo digital– se concentran en tres ejes: sonido como documento/testimonio de la realidad, que representa de manera exacta los acontecimientos recogidos por el sonido; sonido como expresividad, el sonido ambiente es el más expresivo de la diferente tipología sonora, ya que es la mejor manera de representar emociones y sentimientos que no siempre deja entrever la palabra; sonido como ruido, que no sólo se refiere a los efectos sonoros sino a los elementos auditivos que hacen difícil la interpretación correcta.

### **El documento sonoro**

El concepto del documento abarca visiones desde las perspectivas antropológica y cultural, como una extensión o instrumento del ser humano para conservar y transmitir ideas, sentimientos, informaciones; así como desde la visión histórico-jurídica, al ser considerado como una prueba fehaciente de un hecho histórico o jurídico.

Para RODRÍGUEZ (2002), el documento es un mensaje con intención informativa que ha sido incorporado a un soporte que permite su acceso. MARTÍNEZ COMECHE (1998), por su parte, subraya que el rasgo fundamental de todo documento es la transmisión de información. Esas conceptualizaciones pueden considerarse introductorias para la definición del documento sonoro, también denominado fonorregistro, el cual transmite contenidos de información sonora que incorpora o relaciona, uno o todos los elementos del

lenguaje sonoro (voz, efectos de sonido, música y silencio) mediante el uso de recursos técnicos y/o tecnológicos. Entonces, cada documento sonoro es resultado de una grabación sonora, es decir de la fijación de todo tipo de sonidos, de alguna forma material permanente que permite que se puedan escuchar, reproducir, emitir por radio, televisión o comunicarse (IASA, 2005).

En un principio, el término *documento* se refería a la palabra escrita en el sentido de registro de información, pruebas o actividades artísticas o intelectuales. En el siglo XX, su significado se amplió para referirse a la presentación fáctica de acontecimientos, actividades, personas y lugares verdaderos. Una visión más acabada estableció el programa de la UNESCO Memoria del Mundo al sostener que los documentos, con inclusión de los de carácter audiovisual, constan de dos elementos: el contenido informativo y el soporte en que éste se aloja. Los dos tienen la misma importancia.

### **Del documento al archivo sonoro**

La palabra archivo es venida del latín *archivum*, y éste del griego *archeión* que significa principio, origen. En la antigüedad, puede hablarse de la existencia de archivos en las civilizaciones más avanzadas tales como la egipcia, fenicia, persa, siria, mesopotámica y griega.

ALONSO VARELA, NOBLE MOREIRA, y SARAIVA CRUZ (2015) señalan que el surgimiento de los archivos se entrelaza con la historia de la humanidad y con la aparición de la escritura ya que desde esa época se generó documentación registrada, a la vez que se concibió la necesidad de organizarla en orden de conservarla por su valor probatorio. Desde ese período se puede advertir que en la literatura teórica se mencionaba a los archivos como el lugar en el cual se conservaba y custodiaba la documentación, sin aplicarle ningún tratamiento específico. Igualmente se reseñaba que la archivística era una tarea práctica al servicio de la administración. Sobre finales del siglo XVIII y principios del XIX, comienzan a aparecer propuestas metodológicas e intentos de generar leyes generales con relación a la organización de los archivos, así como de los documentos.

Con la Revolución Francesa (1789-1799), comienzan a distinguirse a los archivos como institución, y la actividad archivística se consolidó. Hacia 1898 se publica el Manual de Organización y Descripción de Archivos de los holandeses S. Muller, T. Feith y R. Fruin, que recoge y sistematiza la teoría del historiador francés Natalis de Wally, quien elaboró en 1841 la circular del 24 de abril, en la cual se introduce la noción de “fondo de archivo”. Y de singular importancia para la archivística fue el surgimiento del *Records Management* (gestión de documentos), considerada como una nueva corriente teórica.

Los archivos e instituciones que conservan registros o contenidos de tipo sonoro no aparecieron en una fecha concreta, sino que se fueron materializando en relación con el advenimiento de estos soportes en lugares antes destinados exclusivamente a fondos bibliográficos, durante las últimas décadas del siglo XIX.

Según la Asociación Internacional de Archivos Sonoros y Audiovisuales (IASA, 2021), los archivos de audio desarrollan cuatro tareas básicas: adquisición, documentación, acceso y preservación. Por los influjos de la tecnología, en esos procesos las investigaciones y prácticas relacionadas, incluso, con la inteligencia artificial (IA) están cobrando mucho interés.

## Con el fonógrafo y las fonotecas...

La utilización del fonógrafo condujo a la proliferación de grabaciones de lenguas de pueblos originarios de América, Europa, Asia y África. Y a la aparición de las primeras fonotecas en el ámbito de centros de investigación y de universidades. La Phonogramm Archiv de la Academia de Ciencias y Artes de Viena, fundada en 1899, fue la primera fonoteca del mundo. Posteriormente se formaron fonotecas de investigación por parte de la Sociedad de Antropología, de París, en 1900 y en la Universidad de Zurich en 1908. En ese mismo año, se creó la Phonogrammarchiv de Rusia, como parte de la biblioteca de la Academia Imperial de Ciencias.

El interés y motivación por utilizar la grabación sonora en el trabajo científico de las universidades fue expuesto en El Archivo de la Palabra, impulsado por el profesor Ferdinand Brunot en 1910.

La radio es un componente clave en la producción de archivos sonoros. Según LARSON, BECKERS y SCHLÖGELL (2005), el archivo sonoro de una planta refleja la evolución de sus contenidos a lo largo del tiempo y como tal cumple con la importante función de repositorio como parte de una memoria cultural común.

Es sabido que la radio fue durante muchos años el principal medio generador de documentos sonoros. En la década de los años 30, grabó por vez primera sus contenidos. Dos de los primeros archivos radiofónicos que surgieron fueron el de la BBC (British Broadcasting Company) de Londres creado a partir de 1931 (ROOKS, 2010) y el de Radio Francia Internacional en 1936. A partir de los años 70, además de la música, la voz y los programas de radio, se comenzaron a grabar los sonidos que forman el paisaje sonoro. El antecedente directo de esa práctica, tiene sus raíces en los trabajos de los investigadores Murray Schafer, Hildegard Westerkamp y Barry Truax, de la Universidad Simon Fraser en Canadá. La propuesta formuló la realización de un estudio del paisaje sonoro, en diferentes contextos, con la ayuda de grabaciones de campo y registros acústicos. El paisaje sonoro posibilitó la concepción el contacto con los contextos, y permitió un novedoso uso de lo estético.

## **La alerta: sonidos, patrimonios en peligro**

En la actualidad muchas grabaciones están en riesgo de perderse. Según expertos del área, los archivos digitales de sonido pueden corromperse y los muy utilizados discos compactos regrabables (CD-R) duran pocos años. Uno de los primeros estudios sobre la conservación de las grabaciones sonoras en Estados Unidos, divulgado por la Biblioteca del Congreso, descubrió que muchas de ellas ya se han perdido o que la ciudadanía no tiene acceso a ellas. Se incluye la mayoría de los archivos de la primera década de la radio, de 1925

a 1935. Y los programas con afamados cantantes como Duke Ellington y Bing Crosby, así como las primeras transmisiones deportivas, ya han desaparecido.

Si bien los sonidos en los archivos digitales pueden registrarse y transferirse fácilmente, se exige de una preservación activa. Increíblemente, los viejos formatos analógicos que quedan son físicamente más estables y algunos han sobrevivido unos 150 años más que las actuales grabaciones digitales.

Esos sonidos, hoy más que nunca, son también patrimonio y memoria de culturas que dan cuenta de su devenir.

En esta era, los archivos sonoros poseen usos sociales, políticos, y hasta de entretenimiento. Se han convertido en una fuente de datos central sobre la historia, la cultura y la sociedad, sirviendo de objeto de investigación en diversos campos de la ciencia. Un tremendo desafío al respecto es la falta de un formato estándar para grabar audios. Y cada vez hay más formatos.

El proceso de preservación de un documento sonoro debería iniciarse desde el momento en el que es producido. Los dos componentes básicos a considerar en la preservación digital son el audio digital, para cuya preservación es indispensable que los formatos, resoluciones, soportes y sistemas tecnológicos se adhieran a los estándares aceptados a nivel internacional.

## **Palpando el cambio... nuevas prácticas**

Desde hace unos años, en un intento por recuperar grabaciones históricas y restaurar otras, la Biblioteca del Congreso de Estados Unidos lanzó un plan. PATRICK LOUGHNEY (citado por BBC, 2013), director del centro encargado de la conservación audiovisual, consideró entonces que ese "tipo de preservación es tan importante como salvar edificios, monumentos públicos o parques nacionales, que también son vitales para nuestro patrimonio" (p. 2). Para ello creó una "rockola nacional" (National Jukebox), que digitalizó 10.000 audios de entre 1901 y 1925 y un registro nacional de grabaciones. "Hay una creciente amnesia relacionada con el patrimonio de los sonidos grabados de Estados

Unidos", expresaría. "Ese patrimonio se pierde de una generación a la siguiente porque no está disponible para la investigación pública. Es un problema importante".

En México, en la Fonoteca Nacional se impulsó un proyecto que articuló una audioteca con equipos de cómputo a disposición del público, a través de los que se puede consultar y tener acceso a los archivos sonoros. En la sala de escucha grupal se han realizado círculos de escucha, en los cuales los usuarios pueden participar activamente.

Desde ese país, llegan otras novedades. A través de diversas técnicas de grabación se ha podido realizar un viaje auditivo por la historia de las etnias coras y huicholes, mediante obras reunidas en un disco-libro editado por el Instituto Nacional de Antropología e Historia (INAH), que recoge un siglo de registros musicales y reúne 32 piezas que consignan desde grabaciones realizadas en cilindros de cera hasta las efectuadas con teléfonos móviles.

Otras experiencias incrementan la interacción de los públicos con la música de una manera lúdica y hace uso de las nuevas tecnologías. Además, difunde con todos los internautas la oferta y el contenido musical y educativo de la institución. El proyecto contempla el diseño, desarrollo e implementación de un sistema automatizado de consulta; la instalación de un sistema multimedia para fonoteca; y el diseño, desarrollo e implementación de la biblioteca digital en línea.

En España archivos sonoros registrados en los primeros discos de reproducción de la segunda mitad del siglo XIX se han hecho más accesibles para todo el mundo a través de la plataforma *on line* Biblioteca Digital Hispánica de la Biblioteca Nacional. Se trata de una iniciativa para convertir en archivos digitales de audio una colección de registros audiovisuales que tiene forma de discos perforados que fueron los primeros formatos comerciales de distribución musical. Estos discos realmente no almacenaban el sonido sino los códigos para su reproducción en forma de perforaciones que se podían interpretar en una especie de organillo.

También la Biblioteca Pública de Nueva York tiene una sección de audio donde recoge conversaciones, debates y otros documentos sonoros digitalizados. En el caso de la Biblioteca Pública de los Ángeles también dedica sus podcasts a los adolescentes a quienes hace propuestas de participación como por ejemplo que graben su frase de película favorita, además de recomendaciones realizadas por sus usuarios. La Biblioteca Pública de Denver produce una serie de podcast para niños basados en libros infantiles leídos por sus bibliotecarios. En Europa, la British Library ofrece archivos sonoros de autores en formato podcast en su sección Online Gallery dedicado a autores.

Las aplicaciones<sup>1</sup> para dispositivos móviles son ya una realidad, a través de las cuales las audiencias pueden desarrollar contenidos colaborativos con los archivos. Las funciones esenciales se asocian con promover los servicios, ayudar a consultar los catálogos y permitir compartir en las redes sociales, entre otras prestaciones. Asimismo se incluye material publicitario: una guía de introducción a la institución con imágenes, vídeos e información extra, como números de teléfono o direcciones de correo electrónico a las cuales se llega desde un botón.

Otra práctica interesante es la de la Red Distrital de Bibliotecas Públicas de Bogotá (BibloRed) que construyó la sala Distrito Gráfico de la Biblioteca Pública “Julio Mario Santo Domingo”, experiencia que contó con más de dos mil ejemplares de diversos géneros y formatos en los que prevalecía especialmente la novela gráfica junto a recursos audiovisuales. En experiencias como la anterior, la producción de nuevos formatos discursivos por parte de los usuarios, abre la posibilidad a novedosos espacios a través de los cuales la ciudadanía no solo accede a libros de géneros fantásticos, sino que además experimenta nuevos contenidos gráficos que dinamizan la lectura y la escritura

---

<sup>1</sup> Una aplicación móvil (o app) es una aplicación de software diseñada para funcionar en smartphones, tablets y otros dispositivos móviles.

como la producción transmedia<sup>2</sup>, juegos de rol, serigrafías, xilografías, distintos tipos de ilustraciones y personificaciones.

Otro de los usos del sonido es la confección de mapas acústicos, mapas de ruido o mapas estratégicos de ruido, entendidos como un conjunto de archivos de sonido presentados en formato de mapa interactivo. Esa funcionalidad puede encontrarse con las etiquetas de soundmap, geosonidos, soundscape, cartografía sonora o paisaje sonoro. De forma abierta y colaborativa, permite a los usuarios añadir nuevos sonidos o incluso imágenes asociadas a un punto sonoro.

Para MARTÍ (2012) los mapas sonoros tienen variedad de aplicaciones: etnomusicología, filología, lingüística, sociología, pedagogía, cultura y ocio, entre otras. Adquieren especial importancia en el proceso de creación de comunidades en red al permitir a los usuarios añadir nuevos sonidos, o incluso imágenes, asociadas a un punto sonoro. Por ejemplo, Open Sound New Orleans resultó un proyecto comunitario para que los habitantes de esa ciudad estadounidense documenten sus vidas mediante el sonido. La experiencia pretende hacer accesibles los sonidos y voces auténticos e inéditos de la ciudad, y compartirlos tal y como se escuchan.

SoundCloud es otra de las muestras de los cambios. Ese servicio de alojamiento de audio en la nube<sup>3</sup> se ha ido convirtiendo en una plataforma cada vez más popular, lo que ha llegado a ser vista por muchas discográficas como una amenaza, teniendo en cuenta que actualmente cuenta con más de 175 millones de usuarios únicos mensuales.

---

<sup>2</sup> Proceso narrativo basado en el fraccionamiento intencionado del contenido y su multiplicación a través múltiples plataformas, soportes y canales (*off line* y *on line*), con el fin de que cada medio cuente una parte específica y complementaria de la historia. De esta forma, la comprensión absoluta y el conocimiento profundo de la narración se obtienen cuando se recorren las múltiples plataformas, soportes y canales. Es un proceso que implica interacción por parte del usuario.

<sup>3</sup> La computación en la nube, conocida también como servicios en la nube, informática en la nube, nube de cómputo o nube de conceptos (del inglés *cloud computing*), es un paradigma que permite ofrecer servicios de computación a través de una red, que usualmente es Internet.

Existen además varias alertas ante una especie de futuro cambiante. VERDÚ (2012) asevera que la:

biblioteca digital que recopiló durante años se esfumará con usted. Todos los discos o libros comprados en tiendas electrónicas como Amazon o iTunes (Apple) se perderán en algún sitio del ciberespacio. En realidad, ni siquiera hará falta. Usted nunca los poseyó, simplemente fue arrendatario vitalicio de un servicio de lectura o escucha musical que esas compañías le prestaron a cambio de una cantidad de dinero moleestamente parecida a la que solía pagar por hacerse con un ejemplar físico de esos títulos (p. 2).

En pleno siglo XXI, proliferan empresas de almacenaje como Bluespace para enterrar ahí sin fecha de vuelta todo lo que ya no podemos absorber en nuestra vida doméstica. Y lo mismo sucede con los datos que almacenamos. “Una vez muertos, y ya que nadie más puede tener acceso a ello, lo que queda de nosotros en la red no son más que molestos residuos duplicados infinitamente. Entonces, ¿qué mejor destino que la destrucción?”, se pregunta VERDÚ. Lo físico y tangible cada vez tiene menos relevancia.

## Conclusiones

El trabajo con los archivos sonoros requiere de la articulación de teorías, conceptos y metodologías desde las Ciencias de la Información y la Comunicación.

En la actualidad es imprescindible el diseño de metodologías y programas de preservación digital que permitan proteger los archivos sonoros que son importante patrimonio para lograr que las futuras generaciones tengan acceso a ellos.

Las nuevas prácticas relacionadas con la gestión de archivos sonoros evidencian el influjo de las nuevas tecnologías y de otras tendencias como la inteligencia artificial. En esas nuevas prácticas deberá hacerse énfasis en la participación de las audiencias y en los mecanismos de socialización, estos

últimos contribuirían a que los archivos sonoros lleguen a mayor cantidad de audiencias.

## Bibliografía

- ALONSO VARELA, L.; NOBLE MOREIRA, L.D.; SARAIVA CRUZ, I. (2015). El concepto de paradigma en la Archivística como problema epistemológico. *Palabra Clave*, 4(2), 109-128. <http://sedici.unlp.edu.ar/handle/10915/48169>
- BADÍA VALDÉS, A. T. (2013). La nueva radio: de la producción a la preservación. En línea. <http://www.epoca2.lajiribilla.cu/articulo/6428/la-nueva-radio-de-la-produccion-a-la-preservacion>
- BBC. (2013). EE. UU. rescata su historia sonora. [http://www.bbc.com/mundo/noticias/2013/03/130301\\_cultura\\_eeuu\\_biblioteca\\_congreso\\_preservacion\\_sonidos\\_tsb](http://www.bbc.com/mundo/noticias/2013/03/130301_cultura_eeuu_biblioteca_congreso_preservacion_sonidos_tsb)
- CEBRIÁN HERREROS, M. (1993). *La mediación técnica de la información radiofónica*. Barcelona. España: Editorial Mitre.
- EDMONDSON, R. (2004). *Filosofía y principios de los archivos audiovisuales*. Paris: UNESCO.
- LARSON, M.; BECKERS, TH.; SCHLÖGELL, V. (2005). Structuring and indexing digital archives of radio broadcasters. En: Workshop digitale medianar chive at the 25. Tagung der Gesellschaft für Informatik. [http://www.imk.fhg.de/sixcms/media.php/130/gi2005\\_wdrprojekt.pdf](http://www.imk.fhg.de/sixcms/media.php/130/gi2005_wdrprojekt.pdf)
- FLORES FERNÁNDEZ, C.; GÁTICA MOLINA, C.; GONZÁLEZ CORREA, A.; NÚÑEZ QUINTEROS, T. (2022). Estrategias de preservación digital de archivos sonoros. Revisión sistematizada. *Revista Española de Documentación Científica*, 45(2), e321. <https://doi.org/10.3989/redc.2022.2.1864>
- IASA. (2021). Herramientas de IA para la búsqueda y recuperación de los documentos sonoros digitales con una finalidad educativa. [https://iasa.aviaryplatform.com/collections/1719/collection\\_resources/55627](https://iasa.aviaryplatform.com/collections/1719/collection_resources/55627)
- LINARES COLUMBIÉ, R. (2004). La presencia cognitiva en ciencia de la información y su entorno. En línea. <http://revista.ibict.br/ciinf/article/viewFile/1065/1154>
- MALDEN, S. (2011). Conferencia magistral. En línea. <https://revistas.ucm.es/index.php/CDMU/article/viewFile/38327/37083>

- MALHEIRO DA SILVA, A. (s. f.). La Ciencia de la Información y la transición paradigmática. Material de clase.
- MARTÍ, C. (2012). Mapas Sonoros: un Mashup de sonidos cotidianos. En línea. <https://papelesdemusica.wordpress.com/2012/02/23/mapas-sonoros-un-mashup-de-sonidos-cotidianos/>
- MARTÍNEZ COMECHE, J. A. (1995). *Teoría de la información documental y de las instituciones documentales*. Madrid: Síntesis.
- MENA MUGICA, M. (2015). Conferencia magistral: El cambio de paradigma en el campo de la Archivística. En línea. [http://www.uaeh.edu.mx/xvjornadasarchivisticasrenaies/memorias/conferencias/el\\_cambio\\_de\\_paradigma\\_en\\_el\\_campo\\_de\\_la\\_archivistica.pdf](http://www.uaeh.edu.mx/xvjornadasarchivisticasrenaies/memorias/conferencias/el_cambio_de_paradigma_en_el_campo_de_la_archivistica.pdf)
- PONJUÁN DANTE, G. (2008). Gestión de información: precisiones conceptuales a partir de sus orígenes. *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia* En línea. <http://periodicos.ufpb.br/ojs/index.php/abcib/article/view/5300>
- PRIETO, G. (2016). Mapas sonoros: la cartografía a través del oído. En línea. <http://www.geografiainfinita.com/2016/05/mapas-sonoros-la-cartografia-a-traves-del-oido/>
- RODRÍGUEZ RESÉNDIZ, P. O. (2002). Contexto y desafíos de los archivos sonoros en la era digital. En línea. [http://iibi.unam.mx/publicaciones/294/18%20informacion\\_contextos\\_cambio\\_social\\_perla\\_oliva\\_rodriguez.htm](http://iibi.unam.mx/publicaciones/294/18%20informacion_contextos_cambio_social_perla_oliva_rodriguez.htm)
- RODRÍGUEZ RESÉNDIZ, P. O. (2008). *El archivo sonoro. Fundamentos para la creación de una fonoteca nacional*. Library Outsourcing.
- RODRÍGUEZ RESÉNDIZ, P. O. (2016). La preservación digital sonora. *Investigación Bibliotecológica*, 30(68).
- RODRÍGUEZ RESÉNDIZ, P. O. (2020). *Estado de la preservación digital en los archivos sonoros y audiovisuales de Iberoamérica*. Ecuador: Universidad Andina Simón Bolívar.
- VERDÚ, D. (2012). El fin de la cultura de los objetos. El arrendamiento de los servicios sustituye a la compra y el almacenamiento de bienes culturales. [http://sociedad.elpais.com/sociedad/2012/09/15/actualidad/1347729715\\_451265.html](http://sociedad.elpais.com/sociedad/2012/09/15/actualidad/1347729715_451265.html)
- WOLF, M. (1987). *La investigación de la Comunicación de masas*. Barcelona: Editorial Paidós.

# A importância do povoamento da produção acadêmica de TCCs/TCRs no Arca – Repositório Institucional da Fiocruz

## **CLAUDETE FERNANDES DE QUEIROZ**

Fundação Oswaldo Cruz (Brasil)

[claudete.queiroz@icict.fiocruz.br](mailto:claudete.queiroz@icict.fiocruz.br)

## **LUCIANA DANIELLI DE ARAUJO**

Fundação Oswaldo Cruz (Brasil)

[luciana.danielli@icict.fiocruz.br](mailto:luciana.danielli@icict.fiocruz.br)

## **ÉDER DE ALMEIDA FREYRE**

Fundação Oswaldo Cruz (Brasil)

[eder.freyre@icict.fiocruz.br](mailto:eder.freyre@icict.fiocruz.br)

## **RAPHAEL BELCHIOR RODRIGUES**

Fundação Oswaldo Cruz (Brasil)

[raphael.rodrigues@icict.fiocruz.br](mailto:raphael.rodrigues@icict.fiocruz.br)

## **CATARINA BARRETO MALHEIRO PEREIRA**

Fundação Oswaldo Cruz (Brasil)

[catarina.barreto@icict.fiocruz.br](mailto:catarina.barreto@icict.fiocruz.br)

## **TIAGO MARTINS DA COSTA FERREIRA**

Fundação Oswaldo Cruz (Brasil)

[tiago.ferreira@icict.fiocruz.br](mailto:tiago.ferreira@icict.fiocruz.br)

## **RITA DE CASSIA DA SILVA**

Fundação Oswaldo Cruz (Brasil)

[rita.cassia@icict.fiocruz.br](mailto:rita.cassia@icict.fiocruz.br)

### **VIVIANE SANTOS VEIGA**

Fundação Oswaldo Cruz (Brasil)

[viviane.veiga@iciict.fiocruz.br](mailto:viviane.veiga@iciict.fiocruz.br)

### **PATRÍCIA CARVALHO MENDES**

Fundação Oswaldo Cruz (Brasil)

[patricia.mendes@iciict.fiocruz.br](mailto:patricia.mendes@iciict.fiocruz.br)

### **GLAUCE PEREIRA**

Fundação Oswaldo Cruz (Brasil)

[glauce.pereira@iciict.fiocruz.br](mailto:glauce.pereira@iciict.fiocruz.br)

### **ANA MARIA SOUZA DOS SANTOS**

Fundação Oswaldo Cruz (Brasil)

[ana.souza@iciict.fiocruz.br](mailto:ana.souza@iciict.fiocruz.br)

### **IGOR FALCE DIAS DE LIMA**

Fundação Oswaldo Cruz (Brasil)

[igor.lima@iciict.fiocruz.br](mailto:igor.lima@iciict.fiocruz.br)

### **MARIA DA CONCEIÇÃO CALMON ARRUDA**

Fundação Oswaldo Cruz (Brasil)

[maria.calmon@iciict.fiocruz.br](mailto:maria.calmon@iciict.fiocruz.br)

### **RESUMEN**

O trabalho tem como objetivo apresentar o novo fluxo de depósito das coleções Trabalhos de Conclusão de Curso e de Residências (TCCs/TCRs) dos programas de pós-graduação da Fiocruz no Arca – Repositório Institucional, que será realizado através de autoarquivamento feito pelos próprios alunos, conferido pela secretaria acadêmica e depois aprovado pelo bibliotecário. Foram elaborados

dois tutoriais para ajudar no procedimento do novo fluxo, visando facilitar o trabalho de gestão dos gestores Bibliotecários das Unidades da Instituição e permitir uma melhor recuperação destas coleções. Para a realização deste trabalho foram realizados diversos treinamentos e criados dois tutoriais para facilitar a inserção destas coleções no Repositório. Como resultado, espera-se contribuir para disponibilizar em acesso aberto, como também permitir uma visualização e quantificação destas coleções de forma ágil e eficiente.

#### **PALABRAS CLAVE**

Arca – Repositório Institucional da Fiocruz, Povoamento de repositórios digitais, Acesso aberto ao conhecimento, Repositórios digitais.

Os Repositórios são bases de dados importantes para reunir, disponibilizar e preservar o conteúdo produzido pelas Instituições, permitindo assim, o acesso à informação e aos objetos digitais. É uma fonte de informação confiável, onde o conhecimento científico deve ser certificado por profissionais qualificados e atentos para o preenchimento correto dos metadados. Desta forma, podemos afirmar que os documentos são a essência de uma organização, a memória de uma sociedade, evidenciando a relevância da preservação do patrimônio documental e da visibilidade da informação, e neste sentido:

Os repositórios digitais (RDs) são bases de dados online que reúnem de maneira organizada a produção científica de uma instituição, armazenando arquivos de diversos formatos. Resultam em uma série de benefícios tanto para os pesquisadores quanto para as instituições ou sociedades científicas, proporcionando maior visibilidade aos resultados de pesquisas e possibilitando a preservação da memória científica da instituição (INSTITUTO BRASILEIRO..., 2016).

Salientamos que a criação dos Repositórios foi fundamental para o Movimento em prol do Acesso Aberto que permitiu a consolidação da forma que os autores depositam a sua produção nesses sistemas.

O Movimento de Acesso Aberto ocorreu após a crise dos periódicos científicos nos anos 1980. Para viabilizar o acesso à produção científica foram criadas duas estratégias: a Via Verde (Green Road) e a Via Dourada (Golden Road), que foram definidas na Reunião de Budapest Open Access Initiative (BOAI). A BOAI 2002 também definiu o 1º protocolo de interoperabilidade entre fontes online, o Open Access Initiative-Protocol for Metadata Harvesting (OAI-PMH) (FUNDAÇÃO OSWALDO CRUZ, 2022).

Dentro desse contexto, destacamos o Arca - Repositório Institucional<sup>1</sup> da Fundação Oswaldo Cruz<sup>2</sup>, que reúne, hospeda, preserva, disponibiliza e dá visibilidade à produção intelectual da Instituição. O trabalho de povoamento é imprescindível para o Repositório, pois permite que os documentos depositados possam ser consultados e pesquisados. Em 2014, a Fiocruz publicou a Política de Acesso Aberto ao Conhecimento<sup>3</sup>, que estabeleceu o depósito mandatório dos artigos científicos produzidos em revistas e as dissertações e teses produzidas pelos programas da Instituição. A política preconiza o autoarquivamento dos autores, com etapa de certificação, aprovação e curadoria realizada pelos bibliotecários da Rede de Bibliotecas da Fiocruz, sendo que os Trabalhos Acadêmicos da Instituição possuem um fluxo de depósito pelas bibliotecas, não admitindo o autoarquivamento. Os Trabalhos de Conclusão de Cursos de Residência e de Especialização (TCCs/TCRs)<sup>4</sup>, que fazem parte dos Programas de Pós-Graduação não são considerados coleções

---

<sup>1</sup> O Arca possui 26 coleções entre artigos, dissertações, teses, livros, vídeos, etc. que possuem um fluxo para a certificação e aprovação dos registros. Disponível em [www.arca.fiocruz.br](http://www.arca.fiocruz.br)

<sup>2</sup> Instituição de ciência e tecnologia em saúde vinculada ao Ministério da Saúde (MS) e tem como Missão: “Produzir, disseminar e compartilhar conhecimentos e tecnologias voltados para o fortalecimento e a consolidação do Sistema Único de Saúde (SUS) e que contribuam para a promoção da saúde e da qualidade de vida da população brasileira, para a redução das desigualdades sociais e para a dinâmica nacional de inovação, tendo a defesa do direito à saúde e da cidadania ampla como valores centrais” (FUNDAÇÃO OSWALDO CRUZ, 2022). Disponível em: <https://portal.fiocruz.br/>

<sup>3</sup> Disponível em: [https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/portaria\\_-\\_politica\\_de\\_acesso\\_aberto\\_ao\\_conhecimento\\_na\\_fiocruz.pdf](https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/portaria_-_politica_de_acesso_aberto_ao_conhecimento_na_fiocruz.pdf)

<sup>4</sup> Apenas uma Unidade incluía esta coleção no Repositório, que era o Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT).

mandatórias na Política de Acesso Aberto da Fiocruz e por isso não eram depositados no Arca. Todavia, verificou-se a necessidade desses documentos serem inseridos no Repositório, a partir de uma demanda gerada pelos próprios alunos.



FIGURA 1. Arca - Repositório Institucional da Fiocruz ([www.arca.fiocruz.br](http://www.arca.fiocruz.br))

A Rede de Bibliotecas da Fiocruz, pensando na melhoria dos processos para ampliação do Acesso Aberto e da Ciência Aberta na Instituição, em 2018 estabeleceu um Grupo de Trabalho para se debruçar sobre o tema. Em 2019 foi instituído um subgrupo para refletir sobre novos fluxos dos trabalhos acadêmicos para o Arca, considerando o autoarquivamento pelos alunos da instituição. Tendo em vista a importância da ampliação do acesso à produção científica dos TCCs/TCRs, o Fórum de Coordenadores e a Coordenação Adjunta de Residências - CGE da Vice-presidência de Educação, Informação e Comunicação (VPEIC) criou um Grupo de Trabalho para buscar soluções para o depósito desta tipologia documental no Arca, junto com o GT da Rede de Bibliotecas e a equipe executiva do Arca. Para a discussão do Projeto Piloto foram convidados os gestores da Biblioteca de Saúde Pública (ENSP), da Biblioteca de Manguinhos (ICICT) e da Biblioteca da Saúde da Mulher e da Criança (IFF).

Após a deliberação em algumas reuniões, foi estabelecida algumas orientações e procedimentos para a criação de um novo fluxo de autoarquivamento para os TCCs/TCRs, que deverá ser realizado pelo próprio aluno, validado pela Secretaria Acadêmica e certificado/aprovado pelo profissional Bibliotecário da respectiva Unidade da Fiocruz<sup>5</sup>.

Para ajudar nesse processo, foram criados tutoriais, um para o aluno e outro para a Secretaria Acadêmica, com o objetivo de facilitar no preenchimento dos campos destas coleções no Repositório. Com a finalidade de manter a padronização dos registros, a equipe do Arca organizou diversos treinamentos junto as Secretarias Acadêmicas, professores e discentes pertencentes aos programas de pós-graduação das Unidades Técnico-Científicas da Fiocruz, visando dessa forma, apresentar o novo procedimento, e promover um compromisso com todos para a correta execução do fluxo proposto. Até o momento foram treinadas 65 pessoas, conforme FIGURA 2, sendo que outras Unidades da Fiocruz ainda irão confirmar a data.

UNIDADE	DATA	STATUS	NÚMERO-DE-TREINANDOS
Fiocruz/RO	A-confirmar	A-confirmar	A-confirmar
Fiocruz/ILMD	A-confirmar	A-confirmar	A-confirmar
Fiocruz/Bio-Manguinhos	A-confirmar	A-confirmar	A-confirmar
Fiocruz/ICC	A-confirmar	A-confirmar	A-confirmar
Fiocruz/CE	28/03/2022	Realizado	02
Fiocruz/Farmanguinhos	30/03/2022	Realizado	02
Fiocruz/BSB/GEREB	05/04/2022	Realizado	03
Fiocruz/ICICT	08/04/2022	Realizado	05
Fiocruz/IGM	11/04/2022	Realizado	11
Fiocruz/IAM	13/04/2022	Realizado	03
Fiocruz/MS	13/04/2022	Realizado	04
Fiocruz/INI	18/04/2022	Realizado	06
Fiocruz/ENSP	19/04/2022	Realizado	06
Fiocruz/IRR	25/04/2022	Realizado	04
Fiocruz/EPJIV	26/04/2022	Realizado	04
Fiocruz/IFF	27/04/2022	Realizado	15
Fiocruz/IFF	16/05/2022	Agendada	
Fiocruz/INCQS	17/05/2022	Agendada	

FIGURA 2. Tabela de treinamentos das SECAS Fiocruz

<sup>5</sup> As informações mais detalhadas estão descritas na Nota Técnica 03/2021, disponível em: <https://www.arca.fiocruz.br/handle/icict/51686> (COSER GUTIERREZ et al., 2021).

Em destaque os fluxos, apresentados nas FIGURAS 3, 4 e 5, propondo a inserção dos TCRs/TCCs com as competências de cada seguimento.



FIGURA 3. Competências do Discente



FIGURA 4. Competências da Secretaria Acadêmica (SECA)

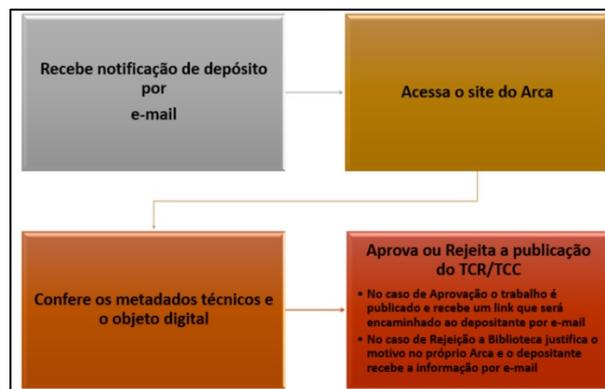


FIGURA 5. Competências do Bibliotecário Gestor do Repositório

Concluimos esse trabalho informando que esse novo fluxo será aplicado a partir do primeiro semestre de 2023 e pretende atender às diversas unidades da Fiocruz com o propósito de ajudar não somente na visibilidade e recuperação dos documentos (TCCs/TCRs), mas também promover o crescimento no povoamento desta tipologia no Repositório Arca. Outro aspecto a considerar é que as unidades poderão acompanhar a produção depositada dos seus alunos (por programa de pós-graduação), e identificar o quantitativo, que atualmente é muito baixo, além de permitir o depósito destas coleções de forma sistêmica e organizada. Acreditamos que ações como essas têm a intenção não só de garantir a confiabilidade e a integridade do Repositório, como permitir e assegurar a salvaguarda, a preservação da memória institucional e dos objetos digitais. O povoamento dessa tipologia é mais uma iniciativa para promover o Autoarquivamento e o Movimento de Acesso Aberto na Fiocruz, através da otimização do novo fluxo informacional dos documentos inseridos e disponibilizados. É preciso que todo esse conhecimento científico produzido pelos alunos dos TCCs/TCRs na área de Saúde Pública, esteja disponível promovendo assim, uma recuperação de forma rápida e eficiente para todos (QUEIROZ *et al.*, 2018).

Durante o trabalho, contamos com a colaboração de profissionais de diferentes unidades da Fiocruz, descritos a seguir: VPEIC/Coordenação adjunta de Residências - Adriana Coser Gutiérrez, Carmen Lucia A. C. Pagotto, Maria Alice Pessanha de Carvalho e Silvana S. Rossi<sup>6</sup>; VPEIC/CINCO: Jaqueline Oliveira e Maria de Fatima Martins; Rede de Bibliotecas: Letícia Ramalho, Adagilson Batista Bispo da Silva, Adriano da Silva e Tatiane Lopes; Ensino/ENSP: Gideon Borges, Mirna Barros Teixeira; Ensino/ICICT: Mel Bonfim e Luciana Martins; Ensino/IFF: Irene Azevedo e Monique Brandão.

---

<sup>6</sup> Que elaboraram a nota técnica 03/2021.

## Bibliografía

COSER GUTIERREZ, ADRIANA et al. (2021). Nota Técnica 03/2021: orientações Fluxo de TCR/TCC das residências para inserção no Repositório Institucional Arca com o apoio da Rede de Bibliotecas da Fiocruz. Rio de Janeiro: FIOCRUZ/VPEIC/CGE. 6 p.

FUNDAÇÃO OSWALDO CRUZ. (2022). Sobre o Arca. Rio de Janeiro, 2022. <https://www.arca.fiocruz.br/terms/sobre.jsp>

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. (2012). Repositórios digitais. Brasília, DF. <https://ibict.br/informacao-para-a-pesquisa/repositorios-digitais>

QUEIROZ, CLAUDETE FERNANDES DE et al. (2018). Desenvolvimento de tutorial para povoamento da coleção de artigos de periódicos no Arca - Repositório Institucional da Fiocruz. In: Seminário Nacional de Bibliotecas Universitárias, Salvador. [Anais...]. Salvador: SNBU. 13 p. Inclui tutorial e comunicação oral. <https://www.arca.fiocruz.br/handle/icict/30715>

# Repensar los Recursos Educativos Abiertos desde los paradigmas de la educación

**MONTSERRAT GARCÍA GUERRERO**

Universidad Autónoma de Zacatecas (México)

[montsegarcia@uaz.edu.mx](mailto:montsegarcia@uaz.edu.mx)

## RESUMEN

El presente trabajo presenta una reflexión teórico-conceptual para mostrar los elementos que pueden recuperarse de diferentes paradigmas educativos para el uso de Recursos Educativos Abiertos (REA), planteado desde la propuesta de apertura de la ciencia. Se parte de las bases epistemológicas de la ciencia para a continuación presentar una reflexión de la posible forma de abordaje de la propuesta de REA desde diferentes paradigmas educativos, con sustento teórico, sobre las diferentes formas de abordar el uso de este tipo de recursos que suponen una apertura no sólo del acceso a la información sino de los procesos de aprendizaje. Se concluye que es viable recuperar elementos de los diferentes paradigmas para enriquecer la práctica educativa frente a una realidad cada vez más demandante de propuestas de acción para la mejora y el cambio y que sólo el conocer el origen y las posturas de los diferentes paradigmas permite rescatar las opciones más viables.

## PALABRAS CLAVE

Recursos Educativos Abiertos; Ciencia Abierta; paradigmas de la educación; sociología de la ciencia.

Las propuestas epistemológicas de autores como Popper, Khun, Merton, Lakatos y Bernal permiten conocer los orígenes de los estudios sobre la ciencia en un nivel amplio; en lo referente al tema de apertura de la ciencia, resultan referentes imprescindibles pues todos ellos se preocuparon por la relación del trabajo científico con la sociedad y pusieron en relieve el tema de las sociedades científicas como núcleos con bases comunes y agendas compartidas, de esta forma el comportamiento de un grupo enfocado a un área de la ciencia resulta tema central en la reflexión sobre sociología de la ciencia. Todos estos pensadores buscaron que la teoría científica/epistemológica estudiara, además de los aspectos formales de la ciencia, el contexto social y político, como quedó representado en el círculo de Viena y tomó mayor importancia cuando a partir de la Guerra Fría las disciplinas científicas fueron apreciadas por sus implicaciones militares y económicas (REISCH, 2009).

La apertura de la ciencia supone la necesidad de que la ciencia llegue a otros niveles diferentes a los que tradicionalmente llegaba; a este respecto, MERTON (1968) ilustra cuales eran los canales de apoyo y control de la ciencia históricamente:

<b>AUSPICIANTE</b>	<b>RECEPTOR</b>
Los Medici	Academia del Cimento
Charles II	The Royal Society of London Observatorio de Greenwich
Luis XIV	Academie des Sciences
Frederick I	Berlin Academy
Pedro el magnifico	Academia de ciencias de St. Petersburgo

TABLA 1. Relaciones de conocimiento y poder

Fuente: Creación propia con base en MERTON (1968)

Estos canales tradicionales funcionaban a manera de patrocinio para el trabajo científico para tener los derechos del conocimiento, las relaciones mostradas en la TABLA 1 fueron transitando al poder de grandes grupos comerciales, donde el ejemplo claro son las empresas farmacéuticas que controlan los descubrimientos médicos, o las empresas de tecnología de la información (innovación tecnológica). Este control de la ciencia llevó a procesos de crítica por parte de los mismos científicos, los gobiernos y las instituciones que financian la investigación y la realidad de que los resultados fueran aprovechados por unos cuantos, sobre todo empresas privadas; de esta necesidad y debate nacen las propuestas de Acceso Abierto (Open Access, OA) que llevaron a que su derivación, aceptada como propuesta de Ciencia Abierta, tenga cada vez mayor incidencia en las prácticas de investigación y de educación, y que existan opciones alternativas como el Buen Conocer en Ecuador.

A partir de la Guerra Fría y con el aumento de apoyo económico para la producción científica, el tema de la medición de la publicación académica tiene dos antecedentes claros: por un lado la necesidad de medir la calidad de las instituciones educativas, lo que derivó en rankings internacionales que a la fecha siguen vigentes y, por otro, la explosión de la cantidad de publicaciones académicas que se realizan, derivada de la inserción de la idea de la ciencia para el desarrollo económico en la agenda internacional, que nace a partir de que se presenta al presidente Roosevelt el informe "Ciencia, la frontera sin fin. Un informe al Presidente en julio de 1945" (BARKSKY, 2014). Estos dos elementos llevan a reconocer un marco teórico específico, enfocado en conceptos como la cuantitativa *ciencia de la ciencia* de BERNAL (1939), *sociología de la ciencia* de MERTON (1968), sobre todo en lo referente al comunismo o comunalismo de la ciencia como uno de los cuatro elementos de su *ethos* de la ciencia, *cienciometría* y el *colegio invisible* de DE SOLLA PRICE (1963), *paradigmas y comunidad científica* de KHUN (1961) y la propuesta de programas de investigación comunes a una comunidad científica de LAKATOS (1989).

Es importante aclarar que para cada uno de los referentes citados en el presente texto no se pretende hacer una revisión amplia de sus aportaciones en todo el ámbito de la epistemología, sino rescatar conceptos y posturas específicas que permiten dar sustento teórico al trabajo, ya que se trata de referentes considerados corriente principal, a la manera de Merton, en temas de abordajes de la ciencia y concepciones sociales de la misma. Esta discusión, que data de por lo menos hace un siglo, da cuenta de la preocupación por los modelos de publicación científica, la necesidad de medir la ciencia, su crecimiento, las estructuras sociales, la política pública o política científica, que llevó a tratar el tema de la apertura de manera cada vez más enriquecida hasta que derivó en propuestas específicas de Acceso Abierto y después de Ciencia Abierta, como ya se dijo.

Robert King Merton, reconocido como el fundador de la sociología de la ciencia moderna, habla de los componentes (*ethos*) de la ciencia, a saber: universalidad, comunismo, desinterés, honestidad y escepticismo organizado y, aunque algunos críticos lo tacharon de ingenuo (BUNGE, 2015), por lo menos desde entonces el tema de la apertura y la propiedad común de los hallazgos (HEES & OSTROM, 2016) ha adquirido un lugar principal en las discusiones sobre ciencia alrededor del mundo. Lo que es importante resaltar es que si bien el acceso abierto nace bajo un esquema basado en conceptos como democratización de la ciencia, conocimiento como bien común y ciencia ciudadana, entre otras propuestas enfocadas a que la ciencia pudiera ser utilizada para resolver las necesidades sociales, las respuestas aceptadas de forma general como propuestas de acceso abierto a la información científica han tomado un camino paralelo a estas propuestas de acercamiento ciudadano, propuestas que pueden considerarse como hegemónicas y que han promovido un discurso de apertura que cambia la matriz productiva más que el acceso a la información científica.

La Ciencia Abierta (CA) es la propuesta contemporánea aceptada a nivel amplio para gestionar la apertura del trabajo científico hacia la sociedad, incluyendo, además del flujo de gestión y publicación del trabajo científico, la

parte educativa; así, el marco que regula el trabajo bajo esta visión debe ir acompañado de propuestas de democratización y vías de acceso efectivas al trabajo científico en un carácter amplio. El uso del concepto inició alrededor del año 2012 y ha sido revisado, discutido y aprobado por diferentes instancias y organismos que le han hecho asumir el papel como paradigma en la forma de manejar la ciencia a nivel mundial (ANGLADA & ABADAL, 2018; VICENTE SAEZ & MARTÍNEZ FUENTES, 2018). Esta propuesta, apoyada por la UNESCO y cuyas recomendaciones se aprobaron en noviembre de 2021, propone una taxonomía que incluye doce elementos, entre ellos, los Recursos Educativos Abiertos (REA) que son el elemento central de análisis propuesto para la evaluación de los posibles abordajes de algunos paradigmas educativos.

Los REA “son materiales de enseñanza, aprendizaje e investigación que utilizan herramientas apropiadas, como las licencias abiertas, para permitir su libre reutilización, su mejora continua y su adaptación por terceros con fines educativos” (UNESCO, 2020, p. 9). Este tipo de recursos han cobrado gran relevancia frente a la pandemia por COVID-19 y se presentan como la oportunidad de ofrecer herramientas de aprendizaje para promover una educación abierta e inclusiva; en consonancia con otras propuestas de apertura de la ciencia. Vale la pena resaltar que esta es una de las propuestas más adelantadas, pues incluso la recomendación de ciencia abierta fue aprobada en 2021 y la de REA en 2019 y existe ya una amplia comunidad de estudiosos y adeptos a este tipo de prácticas, que proponen el uso de recursos compartidos en la educación, lo que transforma la relación profesor-estudiante.

El presente trabajo hace una revisión teórico-conceptual de la forma en que se podrían abordar las propuestas de diferentes paradigmas educativos en la implementación del uso y creación de REA para la educación; se toman como base las propuestas para REA de UNESCO y elementos de los paradigmas conductista, constructivista y sociocrítico para realizar una propuesta de la forma en que se podría abordar el trabajo de creación e implementación de este tipo de recursos en la educación, sobre todo a partir del nivel superior.

## Desarrollo

La propuesta de uso de REA en la educación es cada vez más aceptada por comunidades científicas, lo que supone reconceptualizaciones pedagógicas y teóricas del trabajo frente a grupo; en este sentido resulta interesante recuperar paradigmas educativos y analizar la forma en que se podrían empatar con esta forma de práctica educativa abierta. Los paradigmas que se retoman son el conductista, el constructivista y el sociocrítico, y de todos ellos se retoman algunos conceptos que pueden ser empleados como base de uso de los REA. Se debe aclarar que se usan conceptos o supuestos específicos de las propuestas de cada paradigma, y en ningún momento se pretende mostrar propuestas abarcadoras de lo que implica el trabajo bajo cada paradigma.

### Conductismo

El conductismo supone el uso de estímulos que lleven a una respuesta positiva, y esta postura es criticada por su origen, que toma como base experimentos de laboratorio con distintos tipos de animales. Se propone emplear estrategias que permitan lograr cambios observables a partir de los estímulos; en este sentido se deben utilizar estrategias evaluables, medibles, cuantificables y el receptor/alumno funciona como un destinatario pasivo que está esperando recibir la información que ya viene presentada, al menos en teoría, de una forma que permita tener resultados mínimos del conocimiento que se debe tener para un grado específico de instrucción. Esta propuesta de trabajo educativo es criticada en la actualidad, como lo ilustra SKINNER (1994), al inicio de su libro, pues se encarga de enumerar las críticas a este paradigma, pero también se reconoce que sigue siendo ampliamente utilizada.

Para el caso de los REA se puede pensar en reflejos condicionados y los comportamientos reforzantes (SKINNER, 1994), pues si bien el uso de este tipo de recursos supone que el alumno pueda tener acceso a diferentes opciones de aprendizaje, siempre la elección de temas y la forma de abordarlos responde a propuestas desde el que diseña el REA, propuestas que tienen propósitos de aprendizaje específicos que la mayoría de las veces llevan a una evaluación

para determinar el grado de aprendizaje de los contenidos. Aquí sería importante pensar en el grado de interés del alumno por el tema, la voluntad, la intención y el propósito de búsqueda de aprendizaje. En esta relación de aprendizaje no existe contacto entre alumno y creador de contenido, entonces es el contenido mismo el que tiene que actuar como estímulo o propio estimulante lo que supone un desafío adicional, pues se debe presentar para audiencias amplias y no es posible tomar en cuenta las condiciones generales, por lo que se refuerza la idea de la caja negra que no es posible reconocer de manera amplia.

La propuesta educativa conductista ha sido criticada por ser impositiva, limitada y hasta simplista (LUJÁN & SALAS, 2009), pero sigue representando una forma clara de la relación enseñanza-aprendizaje, de corroborar conocimientos aprendidos que se consideran necesarios y el uso de tecnología cambia la relación, pues se agrega el uso de herramientas y coordinación de variables que puede hacerse aún sin indagar en giros conceptuales que tengan que ver con la psicología psicogenética o social de Vigotsky y Piaget. En este sentido, el enfoque conductista de los REA debe orientarse en contenidos considerados útiles o atractivos y la necesidad de ser presentados de una forma atractiva que permita tener un estímulo de aprendizaje adecuado que se pueda reflejar en mecanismos de evaluación, la idea es planear contenidos para las audiencias más amplias posibles sin pensar en los elementos psicológicos sino en los aprendizajes esperados.

### **Constructivismo**

Para el caso del constructivismo que supone un acercamiento a los procesos psicológicos de quien está aprendiendo con elementos como la epistemología genética de Piaget y la social de Vigotsky, el enfoque de uso y aplicación debe ser otro, de pensar los REA para grupos sociales o públicos específicos, buscando delinear rutas de acceso por características y necesidades, e incluso ir más allá al integrar al alumno en el proceso de creación de su propio aprendizaje, al tener la capacidad de elegir los recursos que mejor convengan, al también darles la posibilidad de fungir como co-

creadores al poder tomar el REA y enriquecerlo desde su experiencia. Esta oportunidad de ser parte del proceso de construcción del recurso permite presentar otra mirada del proceso de aprendizaje, donde las figuras se plantean desde otras posturas, que permiten modificar esquemas de pensamiento existentes.

El proceso de construcción y reconstrucción que supone el constructivismo toma una nueva dimensión bajo los supuestos de la Ciencia Abierta y los REA, pues cualquier recurso puede servir como base de recursos futuros que pueden incluir adiciones del alumno, de procesos grupales y de cualquiera interesado en aprovechar el trabajo para acondicionarlo para otro tipo de estudiante o audiencia; el hecho de que exista apertura no sólo en el acceso sino en el uso y co-creación permite pensar la visión constructivista del uso de REA que supone un progreso continuo y permanente, donde el triangulo instruccional (SERRANO GONZÁLEZ *et al.*, 2011) se enriquece con el uso de tecnologías de la información para poder compartir y enriquecer el trabajo con audiencias más amplias, aunque en este proceso la figura del alumno y el profesor se desdibuja, pues el aprendizaje se vuelve enriquecido desde la colectividad que supone el acceso abierto.

El uso de tecnología en la educación podría llevar a que el alumno pueda construir su propio conocimiento de forma activa e interactiva; para lo anterior es necesario buscar estrategias específicas; como ejemplo, HERNÁNDEZ ORELLANA y LIZAMA LEFNO (2015) proponen el uso de foros o wikis que permitan la interacción entre docentes y alumnos, ya que en plataformas como Moodle y Mooc los contenidos se presentan como inamovibles, pues el guía va colgando las lecciones a las que los alumnos pueden acceder de forma asincrónica, lo que dificulta los procesos de construcción colectiva. En el caso de los REA y los principios de apertura y reutilización se pueden iniciar procesos como los propuestos por PIAGET (1970) de construcción de nuevos conocimientos a partir de los conocimientos previos y se comparten la autoridad y la responsabilidad (SERRANO & PONS, 2011) y los puntos de referencia son más amplios pues el conocimiento sale del aula y se inserta en la comunidad científica.

En resumen, la idea constructivista del uso de REA supone independencia del recurso para ser empleado y enriquecido en procesos de construcción cuasi ilimitados que pueden ir agregando elementos que enriquezcan el contenido de acuerdo al contexto, las necesidades y los descubrimientos e inquietudes de los diferentes usuarios; de esta forma, el proceso rebasa los límites de los triángulos constructivistas propuestos y se puede hablar de una comunidad de aprendizaje más que de una relación de aprendizaje y esto es uno de los principios básicos de apertura que se estarían cumpliendo, bajo la idea de que el conocimiento no sólo es comunal sino que se construye y se debe aprovechar desde y para la comunidad, de aquí la importancia de no usar el constructivismo por moda o slogan o por ser visto como panacea sino como un cambio progresivo de uso de otro tipo de herramientas que aprovechen el uso de tecnologías, como realidad actual, de forma amplia.

### **Teoría sociocrítica**

La teoría crítica nace desde una tradición marxista que busca desde la propia crítica mostrar opciones frente al capitalismo que irrumpía de formas amplia en la realidad mundial en el siglo XIX, bajo esta visión surgen escuelas como la de Frankfurt, con pensadores que buscan caminos para pensar el tema educativo además del económico y el social. En este sentido, destacan pensadores como Adorno, Horkheimer y Habermas, quienes van creando referentes para considerar el entendimiento, el pensamiento, la acción comunicativa, la democracia deliberativa, todos ellos como procesos de entendimiento para la transformación (GÓMEZ DUARTE & PEÑALOZA JIMÉNEZ, 2014) y que redundaron en propuestas como la de Freire, que por un lado son pedagogías del oprimido y por otro de la esperanza y la liberación.

La escuela crítica tiene una larga tradición de pensadores en diversas regiones, y los aportes para la educación han sido muy ricos, siempre con una vena crítica que permite diferenciar la otredad; de esta forma, las propuestas de Estado para la educación y el conocimiento siempre han sido analizadas, entendidas y criticadas como una forma de presión para el cambio, presión de la que resultan propuestas que enriquecen el debate sobre distintos temas

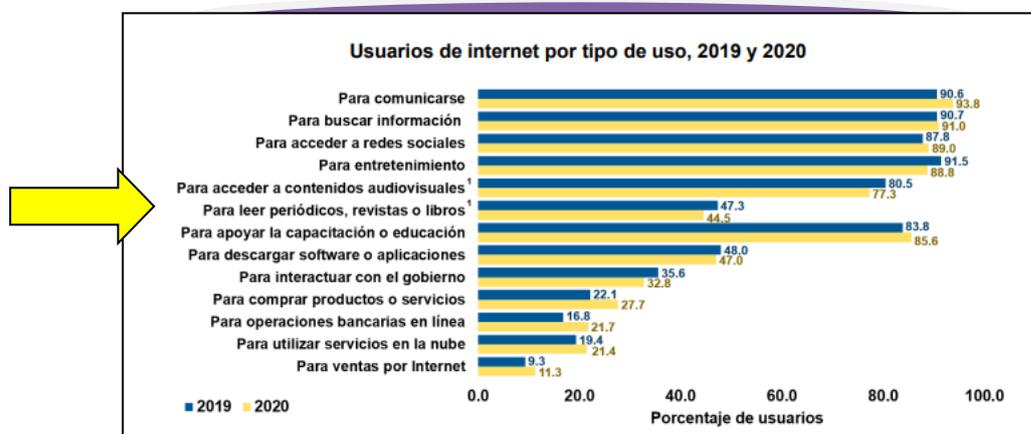
(UNZUETA MORALES, 2009). Para que exista un proceso crítico se debe tener claro el poder hegemónico que privilegia cierto tipo de prácticas y exigencias cada vez más globalizadas como el hablar de competencias, las pruebas estandarizadas, los perfiles de egreso, entre otros.

Para el caso de los REA, al estar apoyados ampliamente por organismos como la UNESCO y al ser usados por instituciones educativas con modelo de empresa se podrían entender como una imposición tecnócrata; desde el trabajo crítico resulta interesante recuperar que la disposición de estos elementos en repositorios de acceso abierto y con opciones de acceso para poblaciones con capacidades especiales los vuelve un tema esencialmente crítico y disruptivo, pues estos recursos se plantean como abiertos e inclusivos y tienen la facilidad de incluir opciones para sordos, débiles visuales y alumnos con otro tipo de necesidades y son además una disrupción en varios temas como la propiedad intelectual, los derechos de autor, pues al ser creaciones colectivas la idea de apertura se va ampliando.

El concepto de REA es ampliamente aceptado por la comunidad educativa actual y de ahí podría verse como imposición, pero si se pudiera llamar de cualquier otra manera lo que realmente importa es la posibilidad de trabajar de una forma inclusiva, reflexiva y comunitaria, donde el uso y generación de recursos responda a la necesidad del colectivo y donde las habilidades de los miembros de este colectivo se puedan aprovechar para enriquecer distintos aspectos. Aunque la crítica siempre lleva a pensar en el acceso a las herramientas tecnológicas y las limitantes, en este sentido, para gran parte de la población, lo que cierto es que la práctica se extiende al uso de aplicaciones para celular. En este sentido, la encuesta de INEGI (2020) ilustra que el 72 % de la población de 6 años o más cuenta con acceso a internet, siendo el celular el principal medio de conexión con 96 %, seguido por la computadora y la televisión, y aunque aún existe una brecha importante entre el sector urbano y el rural: en este último, el acceso está por encima del 50 %.

El uso de los dispositivos para temas educativos resulta un tema nodal para el potencial aprovechamiento de los REA y, en este sentido, vale la pena resaltar

la información de la gráfica de INEGI, en la que se presentan los diferentes usos del internet para los años 2019, 2020, donde el tema educativo aparece en séptimo lugar y con un crecimiento de dos puntos porcentuales entre un año y otro.



GRÁFICA 1. Uso de internet en México

Fuente: INEGI, 2020

La gráfica presentada por INEGI muestra que el uso del internet para fines educativos es ya una realidad nacional y con eso, que es viable pensar en el uso y creación de recursos educativos que promuevan el aprendizaje con herramientas tecnológicas y que aprovechen el uso masivo de los aparatos móviles que hay a nivel global. Donde podría ponerse el punto crítico es en los usos y potenciales peligros del aislamiento por uso de tecnología para temas donde impacta de forma decisiva la socialización, o las formas en que estas tecnologías buscan controlar la vida social y económica de la comunidad.

Lo que es cierto es que los REA pueden servir para ofrecer alternativas educativas paralelas a las tradicionales o institucionalizadas y que estas pueden cambiar sus formas y fondos según los intereses de su(s) creador(es). Bajo esta mirada, esta propuesta de trabajo puede funcionar para el cambio para la esperanza y para pensar en los sueños posibles (FREIRE, 2016), para otorgar mayor claridad para leer el mundo actual y no solo utilizar la tecnología

para el entrenamiento, pues “ser’ en el mundo significa transformar y re transformar el mundo, y no adaptarse a él [...] como participantes activos de la historia” (p. 44). Se puede entonces utilizar la tecnología para ponerse frente al lugar admirado y abstraer para conocer con una conciencia crítica como una forma de empoderamiento y participación en el proceso de enseñanza-aprendizaje.

## Conclusiones

El ejercicio de recuperar la forma en que se puede trabajar desde diferentes paradigmas con un mismo recurso permite reflexionar que el punto de inflexión de la práctica pedagógica no reside en adoptar visiones que parezcan muy cercanas al amplio reconocimiento del otro o a privilegiar el depósito de conocimientos o cambiar la realidad social, sino a buscar la mejor forma de aprovechar los recursos con que se cuenta. El paradigma puede servir como guía que permita recuperar elementos de estudiosos reconocidos en el tema, mas como sucede en el tema de ideología, resultan siempre propuestas que dan ruta pero que resulta imposible llevar a cabo de forma radical, pues se corre el peligro de caer en contradicciones o comprensiones erróneas de lo que se plantea desde una postura específica, ya que las propuestas surgen de realidades históricas que no permanecen estáticas.

Después de este ejercicio se puede observar que el hecho de que un tema sea considerado paradigma no lo exime de ser perfectible, modificable y enriquecible, lo que hace que existan procesos de evolución que permiten responder a las necesidades del momento histórico específico a nivel global, pero también a aquellas del contexto específico donde se realizan los procesos de enseñanza aprendizaje. KHUN (1961) ilustra estos procesos de crisis que permiten llegar a nuevos planteamientos para enfrentar un tema específico; la realidad de la educación actual permite reconocer que es necesario tener la mayor cantidad de herramientas para enfrentar los retos que se presentan en la educación, pues se encuentran problemas de todo tipo, desde pobreza,

limitaciones de aprendizaje, cambios sociales, crisis recurrentes, problemas de salud, imposiciones institucionales, exigencias de indicadores, entre muchos otros que hacen que no se pueda responder siempre desde la misma visión o con las mismas herramientas.

La Ciencia Abierta nace de un discurso amplio que refleja la necesidad de que el conocimiento sea accesible, sin restricciones de pago, para ser aprovechado de forma más amplia, pero también es cierto que aún existen enormes barreras de acceso que hacen que el aprovechamiento de herramientas como los REA aún parezca un tema distante o incipiente, pero el hecho de que existan aún carencias de acceso en la realidad nacional y regional no hace que el uso de estos se vaya a detener y a crear una tendencia de uso a nivel global; por tanto, se debe tener el conocimiento de la forma de elaboración y uso de este tipo de herramientas e ir creando propuestas para enfrentar una realidad que parece se adelanta.

Es posible retomar postulados de cada uno de los paradigmas para atender distintas necesidades con el uso de un mismo recurso, para lo que resulta necesario el análisis del contexto y de la necesidad específica para proponer formas de trabajo desde diferentes ángulos y que el conocimiento de los orígenes de cada propuesta aporte una ruta que permita comprender cuándo aplicar una forma de trabajo u otra. Finalmente los REA, como cualquier otra propuesta de abordaje del proceso educativo, solo servirán en la medida en que puedan ser aceptados por la comunidad y analizados de forma que se presenten diversos trayectos de acción para enfrentar distintos escenarios.

## Bibliografía

- ALVARADO, L. & GARCÍA, M. (2008). Características más relevantes del paradigma socio-crítico: su aplicación en investigaciones de educación ambiental y de enseñanza de las ciencias realizadas en el Doctorado de Educación del Instituto

Pedagógico de Caracas. *Sapiens. Revista Universitaria de Investigación*, 9(2), 187-202. <https://www.redalyc.org/articulo.oa?id=41011837011>

ANGLADA, L. & ABADAL, E. (2018). ¿Qué es la ciencia abierta? *Anuario Think EPI*, 12. <https://recyt.fecyt.es/index.php/ThinkEPI/article/view/thinkepi.2018.43>

BARSKY, O. (2014). *La evaluación de la calidad académica en debate: los rankings internacionales de las universidades y el rol de las revistas científicas*. Ciudad Autónoma de Buenos Aires: Teseo.

BERNAL, J. (1939). *The Social Function of Science*. London: Routledge.

FREIRE, P. (2016). *Pedagogía de los sueños posibles: Por qué docentes y alumnos necesitan reinventarse en cada momento de la historia*. México: Siglo Veintiuno Editores.

GÓMEZ DUARTE, L. & PEÑALOZA JIMÉNEZ, G. (2014). Didáctica y comunicación: aportes de Habermas a la educación. *Praxis & Saber*, 5(9): 13-29.

HERNÁNDEZ ORELLANA, M. & LIZAMA LEFNO, A. (2015) Constructivismo y conectivismo: factor clave para la enseñanza en entornos virtuales. <https://p3.usal.edu.ar/index.php/signos/article/view/3205>

INEGI (2020). Comunicado de prensa Núm. 352/21. 22 de junio de 2021. [https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2021/OtrTemEcon/ENDUTIH\\_2020.pdf](https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2021/OtrTemEcon/ENDUTIH_2020.pdf)

KUHN, T. S. (1961). *La estructura de las revoluciones científicas*. México: Fondo de Cultura Económica.

LAKATOS, I. (1989). *La metodología de los programas de investigación científica*. Madrid: Alianza Editorial.

LUJÁN FERRER, M. & SALAS MADRIZ, F. (2009). Enfoques teóricos y definiciones de la tecnología educativa en el siglo XX. *Revista Electrónica Actualidades Investigativas en Educación*, 9(2): 1-29.

MERTON, R. K. (1968). *Social Theory and Social Structure*. New York: The Free Press

PADILLA BELTRÁN, J. E.; VERA MALDONADO, A.; SILVA CARREÑO, W. H. (2013). La formación del componente pedagógico del docente universitario desde un enfoque sociocrítico". *El Ágora USB*, 13(1): 165-177.

PIAGET, J. (1970). Piaget's theory. In PH Mussen (ed.), *Carmichael's Manual of Child Psychology* (3rd ed, vol 1), Wiley.

POPPER, K. (1959). *La lógica de la investigación científica*. Madrid: Editorial Tecnos.

- REISCH, G. (2009). *Cómo la Guerra Fría transformó la filosofía de la ciencia*. Bernal: Universidad Nacional de Quilmes.
- SERRANO, J. M. & PONS, R. M. (2011). El constructivismo hoy: enfoques constructivistas en educación. *Revista Electrónica de Investigación Educativa*, 13(1). <http://redie.uabc.mx/vol13no1/contenido-serranopons.html>
- SKINNER, B. F. (1994 [1974]). *Sobre el conductismo*. Madrid: Editorial Planeta.
- SOLLA PRICE, D. J. DE (1963). *Little Science, Big Science*. New York: Columbia University Press.
- UNESCO (2020). Directrices para la elaboración de políticas de recursos educativos abiertos. <https://unesdoc.unesco.org/ark:/48223/pf0000373558>
- UNZUETA MORALES, S. (2011). Algunos aportes de la psicología y el paradigma socio crítico a una educación comunitaria crítica y reflexiva. *Revista Integra Educativa*, 4(2), 105-144. [http://www.scielo.org.bo/scielo.php?script=sci\\_arttext&pid=S1997-40432011000200006&lng=es&tlng=es](http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S1997-40432011000200006&lng=es&tlng=es)
- VICENTE SAEZ R. & MARTÍNEZ FUENTES C. (2018), Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research*, 88: 428-436. <doi:10.1016/j.jbusres.2017.12.043>
- WOOTON, B. (1939). Reviewed Work: The Social Function of Science. by J. D. Bernal. *The Economic Journal*, 49(194): 319-321.

# MOOC UTN: una propuesta abierta, masiva y en línea

**PAOLA DELLEPIANE**

Universidad Tecnológica Nacional Facultad Regional BA (Argentina)

[padellepiane@gmail.com](mailto:padellepiane@gmail.com)

## RESUMEN

El concepto de MOOC, cuya traducción en castellano es Curso Abierto Masivo y en Línea, se ha posicionado, en este último tiempo, como centro de discusión, con debates y diversos puntos de vista respecto de las posibilidades, dificultades, críticas y aciertos de esta modalidad educativa. Los cursos masivos en abierto son considerados como un gran fenómeno que afecta la estructura tradicional de la organización universitaria y formativa, y cuyo horizonte próximo resulta impredecible. En este artículo se describe la experiencia de una propuesta de formación en abierto, diseñada y ofrecida por el Centro de Investigación e Innovación Educativa (CIIE) de la Universidad Tecnológica Regional Buenos Aires de Argentina, durante el año 2019. Se desarrollaron tres cursos en formato MOOC para la plataforma MiriadaX (<http://miriadax.net/>). Los cursos se estructuraron por módulos, en los que se proporcionaron contenidos desarrollados en formato video, que son el eje de la propuesta, junto a la interacción y colaboración entre pares. Presentamos, a continuación, la experiencia, con referencias concretas a la manera de encarar la estructura de los cursos desarrollados en formato MOOC, como también los primeros resultados de la implementación en la primera edición durante este año 2020.

## **PALABRAS CLAVE**

MOOC; métodos pedagógicos; diseño didáctico; autoformación.

*MOOC; pedagogical methods; didactic design; self-training.*

## **Introducción**

El Centro de Investigación e Innovación Educativa (CIIE) es una iniciativa de la Secretaría de Ciencia, Tecnología e Innovación Productiva de la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional (UTN). Es un centro de referencia para la investigación e innovación, cuya principal acción es contribuir a la mejora del conocimiento en relación a la tecnología educativa y la enseñanza de la ingeniería.

En esta línea de trabajo, durante el año 2019 se planificó y desarrolló un conjunto de tres cursos, pensados en función de algunas de las características propias de los MOOC, pero manteniendo las líneas de investigación y las características idiosincráticas del CIIE. Estos cursos son:

- Introducción al análisis de malware en Windows (Malware)
- Diseño de experiencia de usuario (Diseño UX)
- Bitcoin y blockchain: una introducción técnica (Bitcoin)

El proyecto de cursos MOOC-UTN se conforma, por un lado, como una vía interesante para ampliar la participación (inclusión) en la educación superior de un público que quizás no tiene posibilidad de acceso a la educación formal; y, por otro, como un medio de promover el aprendizaje durante toda la vida, lo que contribuye al cumplimiento de la responsabilidad social de la universidad.

## **Aspectos generales del MOOC**

Los cursos masivos en abierto son considerados actualmente por muchos investigadores como un gran fenómeno que está afectando la estructura

tradicional de la organización universitaria y formativa, y cuyo horizonte próximo resulta impredecible. Se basan en liberar una acción formativa que se desarrollan en red, es decir, los materiales se encuentran en abierto en un contexto de curso organizado.

Aunque el origen del término se atribuye a Siemens y Downes en 2008, luego del curso Connectivism and Connective Knowledge (CCK08), el concepto de MOOC tiene sus orígenes en la década del 70, en las propuestas de Ivan Illich, quien consideraba que cualquier sistema educativo debía ofrecer formación a todo aquel que quisiera formarse a lo largo de la vida, y alentar a compartir los conocimientos con los demás, de manera pública (ILLICH, 1971). Estas ideas de hace más de cuatro décadas se han podido poner en práctica y potenciar, gracias a las herramientas de comunicación con las que contamos en el mundo digital hoy.

No obstante, la explosión de los MOOC tiene lugar en el 2012<sup>1</sup>, cuyo hito fue el curso abierto sobre inteligencia artificial, impartido en la Universidad de Stanford con más de 160.000 inscripciones de 190 países.

Las principales características que deben tener estos cursos son:

- Tener una estructura orientada al aprendizaje con:
  - Una programación temporal con principio y final.
  - Evaluaciones orientadas a reforzar y demostrar el conocimiento adquirido.
  - Un soporte docente en foros mientras dure la edición.
- Tener un carácter masivo: el número de matriculados es ilimitado, o bien, una cantidad superior a la que podría contarse en un curso presencial.
- Solo es necesario registrarse para poder acceder a los materiales y recursos.

---

<sup>1</sup> Web Inside Higher Education. Audry Watters, <https://www.insidehighered.com/blogs/hack-higher-education/top-ed-tech-trends-2012-moocs>

- No es necesario ser estudiante de la universidad que lo imparte para poder participar.

## Características de los MOOC

El potencial de los MOOC reside en su capacidad para el desarrollo de competencias transversales, que implican la organización, planificación y gestión de la información y el tiempo. No obstante, se hará a continuación referencia a dos aspectos relevantes que implican el concepto de MOOC que toman protagonismo en el análisis colectivo actual, y que son: *masivo* y *abierto*.

El carácter masivo de los MOOC es una de las características que los diferencian de otras actividades y cursos en línea, y es también uno de los aspectos que continúan en el debate, principalmente vinculado a la calidad de la propuesta formativa y el acompañamiento e interacción con el docente o tutor. El interrogante que aún continúa es ¿hasta qué punto un curso con tantos alumnos tiene calidad educativa? (PRENDES & SÁNCHEZ, 2014).

El concepto de abierto refiere a que los usuarios se pueden registrar sin ningún tipo de condicionamiento previo. Así, el cambio que suponen los MOOC respecto a experiencias educativas de comienzo de siglo, como objetos de aprendizajes, repositorios digitales abiertos, OER, OCW, se refiere principalmente al hecho de poder realizar de manera totalmente gratuita y abierta un curso en línea. Ya no nos referimos al acceso de recursos educativos abiertos, sino a poder participar en un curso con actividades, recursos, tests y pruebas de evaluación, y en unas fechas concretas (PRENDES & SÁNCHEZ, 2014).

## Estructura de un MOOC en MiriadaX

Basándonos en estos principios y siguiendo las especificaciones de Telefónica Educación Digital (TED), hemos desarrollado los cursos abiertos en la plataforma virtual MiriadaX (<http://miriadax.net>).

MiríadaX nació en 2013 con la finalidad de fomentar la difusión del conocimiento en abierto en el espacio Iberoamericano de educación superior. Junto a las universidades participantes, se ofrecen cursos de diversas temáticas, diseñados para poder seguirlos al ritmo de cada uno y en los que se puede aprender desde cero sobre alguna temática o profundizar en ellas.

La plataforma MiriadaX es un entorno virtual de enseñanza que cuenta con la mayoría de herramientas y principios de cualquier LMS convencional. Incluye un sistema de gestión de usuarios que implica su registro previo para poder acceder al espacio correspondiente al curso de interés.

Siguiendo con esta línea, el proceso de aprendizaje se desarrolla en forma simultánea en un plano individual y en otro social, a través de las interacciones y discusiones colaborativas que puedan generarse en los foros del aula virtual.

El contenido del curso debe desarrollarse principalmente con recursos audiovisuales que contemplen dos horas promedio mínimo de video para un curso de cuatro semanas, junto a un apoyo textual que esquematice el contenido desarrollado en los videos.

Los cursos en MiriadaX proponen una estructura y parámetros mínimos que deben cumplirse, para una duración mínima de cuatro semanas y máxima de doce semanas, y que pueden sintetizarse en la siguiente estructura:

- Ficha introductoria
- Guía didáctica
- Módulos de contenidos

## **Experiencia de la universidad en el desarrollo de MOOC**

Durante la primera mitad del año 2020, se impartieron los tres cursos con más de 6.400 inscriptos. Para el diseño de los cursos y producción de materiales, se pensó un proceso en tres etapas:

**Etapa 1 – Planificación:** se organizó el curso, se definieron los módulos que conformarían cada uno de los cursos y la secuencia de actividades. Para ello, el equipo se reunió con cada docente para ajustar las definiciones.

**Etapa 2 – Diseño y producción de videos:** se realizaron las grabaciones de los videos para su posterior edición, considerando los criterios definidos por MiriadaX.

**Etapa 3 – Implementación:** se montaron todos los cursos en la plataforma para su validación. Finalmente, se definieron las fechas de inicio y se procedió a la apertura de la inscripción.

### Algunos resultados de la implementación

En el siguiente esquema se presenta una síntesis estadística de inicio y finalización de los cursos:



IMAGEN 1. Inscripciones y tasa de finalización de los tres MOOC. Elaboración propia

Se realizó un cuestionario de cierre y valoración para conocer el perfil de los participantes que iniciaron cada uno de los cursos, así como conocer su

opinión y sugerencias de mejora. Respondieron 138 (Malware), 83 (Diseño UX) y 118 (Bitcoin).

A continuación se presentan algunos de los resultados de los atributos consultados:

- En los cursos de Malware y Bitcoin, el 79 % de los que realizaron el curso son varones. En tanto, en el curso de Diseño UX, la proporción resultó más distribuida, siendo el 51 % varones.
- En los tres cursos, el 70 % de la población participante, se encuentra en el rango de edad entre 25 y 55 años. Para el caso del curso de Bitcoin, el rango de mayor proporción corresponde al rango de edad de 46 a 55 años (25 %). En tanto, para los cursos de Malware y Diseño UX, el rango de mayor porcentaje (33 %) corresponde al intervalo de 36 a 45 años.
- Un poco más del 70 % de los participantes que realizaron los cursos de Bitcoin y Malware finalizaron estudios universitario y de posgrado. En tanto ese porcentaje es de 50 % en el caso del curso de Diseño UX.
- En los 3 cursos, el 80 % no se desempeña como docente.
- En los cursos de Malware y Diseño UX la motivación a realizar el curso fue “Adquirir nuevas habilidades”, en tanto para el curso de Bitcoin, resultó mayoritariamente el “Interés personal”.
- Respecto a la presentación de los contenidos y claridad en la estructura, en los tres cursos el 80 % de los participantes lo valoró entre Excelente y Muy buena.

En relación con la metodología, algunos comentarios refirieron a solicitar más cantidad de bibliografía complementaria. Las tasas de abandono resultaron dentro de lo esperable, no obstante, el 25 % promedio entre los tres cursos, manifestó que era el primer MOOC que realizaban.

## Conclusiones

Sin dudas, se trata de una nueva forma de concebir la enseñanza, en la que se diluyen las barreras temporoespaciales, al tiempo que se transforman los roles y se modifican las intencionalidades de enseñanza de los docentes.

La gratuidad y masividad son dos conceptos que diferencian un MOOC de otro tipo de formación e-learning tradicional. Sin embargo, para que este movimiento siga avanzando, es necesaria una reconceptualización y readaptación que genere un modelo pedagógico y didáctico sostenible en el tiempo, fundamentalmente en lo que hace al rol del estudiante y a las prácticas docentes.

El desarrollo de los tres cursos implicó un año de trabajo interdisciplinario con los docentes referentes y el equipo tecnológico para la elaboración de los materiales y la producción de los más de sesenta videos o píldoras de aprendizaje resultantes. No obstante, consideramos que este modelo puede ser una gran oportunidad para aprovechar las múltiples posibilidades que nos ofrecen las tecnologías para generar aprendizaje de calidad, con un diseño pedagógico y colaborativo elaborado críticamente.

Adicionalmente, los MOOC resaltan la importancia de la formación en red, y con ello la necesidad de repensar la estrategia educativa para fomentar la apertura de la universidad a la sociedad. Pensando en prospectiva, los MOOC pueden funcionar en el engranaje universitario como una pieza más del sistema, considerando una oportunidad más que una amenaza. En este contexto, las universidades tendrían que desarrollar mecanismos para el reconocimiento de créditos en titulaciones formales en un nuevo modelo de educación a distancia y formación virtual.

## Bibliografía

- CABERO-ALMENARA, J.; LLORENTE, M. C. (2017). Los MOOC: encontrando su camino. *@tic, Revista d'Innovació Educativa*, 18, 24-30. [https://grupotecnologiaeducativa.es/images/bibliovir/Los\\_MOOCencontrando\\_su\\_camino.pdf](https://grupotecnologiaeducativa.es/images/bibliovir/Los_MOOCencontrando_su_camino.pdf)
- ILLICH, I. (1971). *Deschooling Society*. Marion Boyars. London and New York.
- PRENDES ESPINOSA, M.; SÁNCHEZ VERA, M. (2014). Arquímedes y la tecnología educativa: un análisis crítico en torno a los MOOC. *Revista Interuniversitaria de Formación del Profesorado*, 28(1), 29-49. <https://www.redalyc.org/pdf/274/27431190003.pdf>
- SIEMENS, G. (2004). Conectivismo: Una teoría de aprendizaje para la era digital. <http://es.scribd.com/doc/201419/Conectivismo-una-teoria-del-aprendizaje-para-la-era-digital>
- VIZOSO MARTÍN, C. (2013). ¿Serán los COMA (MOOC) el futuro del e-learning y el punto de inflexión del sistema educativo actual? *Revista Intenciones. Tendencia en Innovación e Intervención Social*, 5.

Eje temático:  
Datos abiertos

[Ver bloque de ponencias en You Tube](#)

# Gestión de los datos para la investigación en los repositorios del área de ciencias de la tierra de la UNAM

## SAÚL ARMENDÁRIZ SÁNCHEZ

Instituto de Geofísica, Biblioteca Conjunta de Ciencias de la Tierra (UNAM) (México)

[asaul@atmosfera.unam.mx](mailto:asaul@atmosfera.unam.mx)

## MINERVA CASTRO ESCAMILLA

Instituto de Geofísica, Biblioteca Conjunta de Ciencias de la Tierra (UNAM) (México)

[minerva@igeofisica.unam.mx](mailto:minerva@igeofisica.unam.mx)

## RESUMEN

La Universidad Nacional Autónoma de México (UNAM) cuenta con seis entidades académicas dedicadas a la investigación que generan importantes volúmenes de datos para la investigación con base en diversos proyectos y considerando distintos equipos que se encargan de generarlos y procesarlos para ponerlos a disposición de los usuarios, de una manera abierta a través de sus Repositorios Institucionales, combinando ahora los datos con literatura científica para ampliar su posibilidad de visibilidad, impacto, consulta y recuperación, integrándose con ello al tema de los datos abiertos en México. Con base en ello, el presente documento se encamina a mostrar el proceso de gestión de los repositorios generados por el Instituto de Geofísica de la UNAM (SSN, Geotermia y Clima Espacial) estableciendo desde su planeación, estructuración de metadatos y liberación. Estas tres etapas son fundamentales para que un repositorio de datos para la investigación se convierta en una herramienta clave para una especialidad, que en nuestro caso son las ciencias de la tierra.

De igual manera la descripción de las etapas llevadas a cabo en su diseño, los procesos y el software empleado han sido dos de los elementos que permitieron su desarrollo, para lo cual se expone cada uno de ellos para entender la estructura del trabajo. Por ello, el objetivo principal del trabajo es mostrar todas estas actividades bajo proyectos reales y que se encuentran en operación en estos momentos, apoyando la ciencia no solo nacional sino también mundial.

#### **PALABRAS CLAVE**

Gestión de datos para la investigación; ciencias de la tierra; repositorios institucionales; ciencia abierta; literatura científica; marcaje de datos; conversión y migración de datos.

*Data management for research; Earth sciences; institutional repositories; open science; scientific literature; data marking; data conversion and migration.*

## **Los datos en Ciencias de la Tierra en México**

El aumento de la actividad científica produce cambios profundos a nivel mundial con sofisticados avances en Tecnologías de Información y Comunicación y en el incremento de la inversión, tanto pública como privada. El avance en áreas como las Ciencias de la Tierra, el Mar y la Atmósfera, así como en la Astronomía, la Física o las Ciencias Biomédicas y de la Vida está permitiendo la obtención y la recolección de datos, que, además, crecen de manera exponencial. Como consecuencia, tanto investigadores como organizaciones dedicadas a la investigación se encuentran cada vez con más dificultades para gestionar, analizar e interpretar esta producción de datos en grandes volúmenes.

Al hablar del desarrollo de la ciencia en México, y sobre todo al referirnos a las Ciencias de la Tierra en donde su producción científica nacional es de aproximadamente 33.991<sup>1</sup> documentos indizados (que representa el 47,2 % de la producción que se realiza en América Latina), cuenta con difícil acceso a través de plataformas o estructuras propias o locales, aunándole a ello si el acceso es por suscripción, registro, pago o que sean financiados con recursos públicos.

Esta área de la ciencia ha crecido rápidamente en la última década, convirtiéndose en materia prima con un parámetro de desarrollo económico entre los países, en donde se invierten grandes cantidades de recursos para su desarrollo por ser prioritaria para la sociedad. Un problema al que se enfrenta nuestro país es la falta de difusión de información científica y tecnológica que se deriva de una serie de investigaciones que se llevan a cabo tanto por entidades públicas, privadas y de investigación.

Bajo este panorama y debido al gran impacto de los diversos fenómenos naturales que han golpeado a México, así como el aumento de datos para la investigación e información académica, se considera necesario facilitar a la sociedad herramientas que permitan la consulta de estos datos de manera libre, oportuna e intuitiva para entender los cambios que sufre el orbe desde el punto de vista de las Ciencias de la Tierra. Temas como el calentamiento global, desarrollo sustentable, cambio climático, variación de oleaje en los mares, movimientos sísmicos, estados geotérmicos en el país, así como de clima espacial, entre otros, son estudiados por los científicos, y en ocasiones como resultado desarrollan infraestructuras tecnológicas en beneficio de la humanidad y las propias ciencias de la tierra.

---

<sup>1</sup> Resultados de la búsqueda de información en el Sistema de Administración del Conocimiento Web of Science, sobre la producción científica en Ciencias de la Tierra en México. Mayo 11, 2022.

Ante la producción masiva de datos, se requiere facilitar el acceso, el intercambio, almacenamiento, diseminación, así como la difusión de información especializada en Ciencias de la Tierra, mediante herramientas funcionales que permitan a los investigadores y a la sociedad consultar de manera abierta sus líneas de acción y llevar a cabo un mejor estudio de las variantes que se presentan en México y en otras partes del mundo, a fin de contar con información retrospectiva, actual y en tiempo real, producida en diversas entidades especializadas en Ciencias de la Tierra.

Estas herramientas deben apoyen el desarrollo y la preservación de la ciencia por medio de una serie de procesos y flujos de trabajo de la información (datos), generados por organismos, servicios nacionales, oficiales y de investigación, que abordan temas de prioridad y seguridad nacional, como los desastres naturales, cambio climático y temblores de tierra por ejemplo, los cuales se han incrementado en los últimos años y cada vez son más cercanos en tiempo.

Si consideramos que México cuenta con 28.578<sup>2</sup> investigadores, según el Portal de Datos del Sistema Nacional de Investigadores (SNI), y 27,09 % de la producción científica en Ciencias de la Tierra es generada por la UNAM, se necesitan mecanismos y recursos que puedan apoyar la difusión de toda esta producción mediante instrumentos que permitan el manejo, estudio, análisis y uso de los datos, lo que proporcionará un fortalecimiento de la información, la cual puede prevenirnos de cualquier eventualidad en un desastre natural.

## **La literatura científica en apoyo a la ciencia abierta**

En los últimos años, el movimiento Open Access (OA) o Acceso Abierto ha tomado fuerza entre las instituciones académicas y científicas, por lo que este

---

<sup>2</sup> Conacyt. Sistema Nacional de Investigadores (SIN). Tabla de Investigadores Vigentes 2018 por Grado, Nivel, Género, Adscripción, Entidad Federativa y Área de Conocimiento. Disponible en <https://datos.gob.mx/busca/dataset/sistema-nacional-de-investigadores>.

concepto se relaciona directamente a publicaciones de investigación, que buscan favorecer y promover el acceso libre y sin restricciones a los trabajos generados por la comunidad científica a nivel mundial. El Acceso Abierto a la literatura académica y documentación de investigación es vital para la comprensión de nuestro mundo y para la búsqueda de soluciones a los desafíos globales, en especial para reducir la desigualdad de la información, garantizando la integridad del sistema de comunicación académica, asegurando que toda investigación estará disponible a perpetuidad.

En cuanto al panorama actual del estado que guarda el Acceso Abierto, se encuentran inmersos 158<sup>3</sup> países, divididos en cinco regiones. Revisando estadísticas e informes, pocas naciones han tenido éxito, por su entorno y los factores que los rodean en la implementación; en otros casos, el avance ha sido bajo o se encuentran en sus primeras etapas de desarrollo.

Existen iniciativas regionales de Acceso Abierto, que ofrecen a los autores y usuarios finales publicaciones académicas y científicas de texto completo, en portales de revistas multidisciplinarios revisadas por pares. Plataformas patrocinadas principalmente por fondos públicos, buscando la visibilidad internacional, indicadores bibliométricos para una mejor calidad de los procesos editoriales, tales como DOAJ y SciELO. También hay portales como el de las revistas en Brasil –como SEER, a nivel institucional de la Universidad de Sao Paulo (USP)–. En México, el portal de revistas académicas y arbitradas de la Universidad Nacional Autónoma de México (UNAM), que tiene como objetivo facilitar la búsqueda y consulta de las revistas e incrementar su visibilidad, presencia e impacto. Por otro lado, la Universidad de Chile tiene un significativo número de revistas de Acceso Abierto, por dar algunos ejemplos.

Por ello, la literatura científica en apoyo a la ciencia abierta contribuye a mejorar la metodología y calidad de los procesos editoriales, la visibilidad de las revistas académicas editadas en América Latina y el crecimiento de la

---

<sup>3</sup> The Global Open Access Portal (GOAP), UNESCO. Disponible en <http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/>.

adopción del Acceso Abierto. Y, más recientemente, los Repositorios Institucionales de centros y universidades públicas y privadas. Si nos remontamos al año 2002, se ha fortalecido en gran medida la ciencia y su compartición, así como el reúso de datos por parte de los investigadores a nivel mundial. México, en el 2014, crea una iniciativa que logra integrar el Capítulo 10 de Ciencia y Tecnología sobre Acceso Abierto a la información científica y tecnológica, creando el Programa de Repositorios del Consejo Nacional de Ciencia y Tecnología (Conacyt)<sup>4</sup>, implementándose el Repositorio Nacional (RN), y a finales de ese mismo año se crean los Lineamientos Generales para el Repositorio Nacional (RN) y los Repositorios Instituciones (RI)<sup>5</sup>.

Poco después, en 2019, se reunieron los representantes de la African Open Science Platform, Amelica, Coalitions, OA2020 y SciELO, cinco de las principales iniciativas mundiales de Acceso Abierto en la reunión anual de la Declaración de São Paulo, con el fin común de hacer que el conocimiento esté disponible y sea accesible en cualquier lugar, pues este ayuda a resolver los desafíos de la humanidad.

## Proceso de datos

Este proceso de recopilación de información y análisis de los datos en las Ciencias de la Tierra resultan esenciales para la comprensión del entorno natural, pues cambian a cada momento (tiempo). A diferencia de otros campos científicos que están muy alejados de la vida de la mayoría de las personas. En esta área se tratan los procesos que influyen todos los días en el mundo, como las tasas de lluvia, tendencias en las pesquerías marinas, probabilidades de terremotos, potenciales de erupción volcánica, entre tantas otras.

---

<sup>4</sup> Conacyt (2022). Repositorio Nacional. <https://www.repositorionacionalcti.mx/>

<sup>5</sup> Conacyt (2017). Lineamientos generales de Ciencia Abierta: Conacyt. México.

Es por eso, que consideramos y partimos de la declaración de la American Geophysical Union (AGU),<sup>6 7</sup> en donde convierte a los datos en un patrimonio mundial, que deben ser conservados a largo plazo para futuro uso, pues ayudarán a los científicos a comprender los sistemas terrestre, planetario y heliofísico. Y se tendrá la obligación a ponerlos a disposición de la comunidad científica y del público tan pronto como sea posible, debiendo ser accesibles, en formatos utilizables, con suficiente documentación legible que permita su reutilización.

La adhesión a esta política fomenta los avances científicos y produce beneficios económicos que repercuten en la toma de decisiones, a fin de mejorar la seguridad pública y el bienestar social, mismos que contribuyen a la seguridad nacional y mundial al mantener a la sociedad más informada.

Asimismo, busca ser lo más abierta posible para maximizar el rendimiento científico, económico y social de la obtención de los datos, ya sea con fondos públicos y/o privados. Solo existirán situaciones limitadas de acceso por cuestiones éticas, pero no por propiedad. Estas deben ser justificadas a un nivel de protección de seguridad, la confidencialidad para así poder evitar daños a entidades protegidas. Lo importante es documentar las tendencias y la producción de los datos a largo plazo para comprender muchos fenómenos naturales, debido a que el estado de los sistemas naturales nunca se repite, las pérdidas de datos o las oportunidades de recopilación no pueden ser corregidas.

La AGU es clara tanto en su postura, considerando el valor de los datos que crece con el tiempo, llevando a cabo ventajas en la curación de datos. Por lo tanto, los datos y la documentación asociada deben depositarse en un

---

<sup>6</sup> La American Geophysical Union, establecida en 1919 por el Consejo Nacional de Investigación, asociación científica internacional sin fines de lucro –con 60.000 miembros en 137 países– crea su Declaración.

<sup>7</sup> Declaración adoptada por la American Geophysical Union el 29 de mayo de 1997; y fue refirmado en mayo de 2001, mayo de 2005, mayo de 2006 y finalmente revisado y refirmado en mayo de 2009, febrero de 2012 y septiembre de 2015.

repositorio de la especialidad, siendo parte integral de los programas operativos y de la investigación, por lo que la comunidad científica tiene un amplio conjunto de datos para la realización de sus investigaciones en beneficio de la sociedad y de la ciencia.

Con el desarrollo de iniciativas internacionales de investigación a gran escala, el apoyo a los datos tomó una variedad de formas. Se establecieron una serie de instalaciones de datos, por ejemplo, los Centros de Datos Mundiales y la Federación de Servicios de Análisis de Datos Astronómicos y Geofísicos<sup>8</sup>; surgió una visión revolucionaria en el área científica de las Ciencias de la Tierra, centrando la atención de los geocientíficos en conjunto en métodos científicos, mediciones y datos.

En cuanto a recursos nacionales mediante infraestructuras tecnológicas para los investigadores. Estos centros se convirtieron en depósitos de datos junto con otras instituciones de preservación, como archivos con colecciones de fotos y manuscritos, museos con objetos físicos y bibliotecas con libros y revistas.

Actualmente, hay debates científicos sobre la producción de grandes datos y de infraestructuras que se centran en investigaciones, reavivando el interés social y cultural, transformándose en nuevas líneas de investigación y de conocimiento.

## La gestión de contenidos

Para trabajar en la gestión de estos contenidos, tanto la información como los datos ambos deben ser accesibles, interoperables y reutilizables no sólo para las personas, sino también para las máquinas con identificadores, incluyendo metadatos únicos y persistentes, apropiados para ayudar al

---

<sup>8</sup> Surge en 1957 el Año Geofísico Internacional, en inglés International Geophysical Year (IGY), el cual ha evolucionado con una visión revolucionaria en el área científica de las Ciencias de la Tierra.

descubrimiento y ser citados de manera equivalente a otros resultados académicos, mediante un protocolo estándar basado en la web, proporcionando la información de procedencia y el licenciamiento de uso y sean interoperables técnica y semánticamente, así como reutilizables.

Los datos también deberán estar bien conservados, ser accesibles y vinculados de forma segura a las publicaciones asociadas, recursos o facilitar el reuso.

Existen estándares para el uso y tratamiento de los datos y se deben contar con las siguientes características:

1. Para publicar los datos en cualquier formato se debe contar con una licencia abierta, con finalidad de que el público que los encuentre tenga el beneficio de acceder libremente a la información bajo una licencia de uso.
2. También los datos que sean publicados bajo una estructura con licencia abierta deberán estar estructurados en un formato legible por máquina, por lo que el usuario podrá reutilizar los datos publicados y hacer su propio análisis.
3. Se deberán usar formatos no propietarios, es decir, datos abiertos, estructurados, y sin propietarios, con el objetivo de que ellos se consoliden y ofrezcan potencial para que su reutilización sea fácil, productiva, además de que genera impacto.
4. Se deben emplear URI con la finalidad de integrarse con otras fuentes de datos y aplicaciones, pasando de datos abiertos a datos vinculados, por lo que se facilita su identificación y localización mediante “identificador uniforme de recursos” o URI.
5. Finalmente los datos deberán enviar a otros datos para proveer contexto. Por ello, se vincularán y se relacionaran con otros datos de información relevante, la fusión de estas dos o más fuentes de información pone todo este conocimiento en un nuevo contexto, creando valor adicional.

Es importante mencionar que para la gestión de los datos para la investigación en los repositorios del área de Ciencias de la Tierra, se debe considerar además un elemento esencial de apoyo, refiriéndonos a las publicaciones académicas debido a que es una fuente utilizada por la comunidad científica para dar a conocer los avances de sus investigaciones, siendo por lo regular, que los datos se complementan, respaldan y/o vinculan con la publicación (artículo, informe, libro, capítulo, tesis, etc.) o un proyecto ofreciendo como resultado su presentación.

Con el fin de llevar a cabo el repositorio para manejar y hospedar datos, se pueden considerar una serie de buenas prácticas en su desarrollo, dentro de los cuales podemos destacar:

- Citación y procedencia en publicaciones académicas –cuando se considere necesario–. Cuando los investigadores utilizan datos creados por otros, estos deben ser citados con referencia a su autor, a su procedencia y a un identificador digital permanente.
- Interoperabilidad en la mayor medida posible. Tanto los datos de investigación como los metadatos que permiten la evaluación y reutilización, todo ello con el objetivo de que el intercambio y el enlace de plataformas e información permita un flujo transparente para el usuario.
- Reutilización no restrictiva. Los datos para investigación que están en el dominio público deben ser etiquetados como reutilizables por medio de una renuncia a los derechos o una licencia no restrictiva, mediante la cual se deje en claro que los datos pueden ser reutilizados sin mayor requisito que el reconocimiento al autor.
- La información debe llevar el dataset o el dato del lugar donde se hospedarán –que se incluirá en el repositorio dentro del archivo llamado *readme file* con la siguiente información y descripción:

- I. Proporcionar la información sobre los datos y/o el set de datos para que sea correctamente interpretado tanto por personas como máquinas y se puedan realizar procesos con otras investigaciones.
  - II. Cuando el registro (unidad de descripción) se compone de varios sets de datos, se debe vincular a un archivo que se llame *readmefile* (archivo que contendrá la descripción de los datos), de tal manera que pueda ser asociado al set de datos y se especifique en el archivo de formato de fácil consulta como csv o txt.
  - III. Debe incluir una breve descripción de los datos según sea el caso: contacto del investigador principal, fecha de recolecta de datos y de creación del set de datos; información geográfica de los datos; metodología y enlace a publicaciones y otra documentación; unidades de medida, protocolos, abreviaciones, códigos, símbolos asociados a los datos; licencia de uso; citación recomendada.
- Capacidad de vinculación de los datos abiertos, siempre que sea posible estar vinculados con otros datos basados en su contenido y contexto, con el fin de maximizar el valor semántico.

## Los principios rectores de FAIR para la gestión y la administración de los datos científicos

La necesidad urgente de mejorar la infraestructura que respalda a la reutilización de datos académicos y el grupo de personas interesadas –que representan a la academia, la industria, las agencias de financiamiento y los editores académicos–, se unieron para diseñar y respaldar los Principios (Hodson, 2021) que referimos, con la intención de utilizarlos como una guía para aquellos interesados que deseen mejorar la reutilización de sus datos. A diferencia de las iniciativas de pares –que se centran en especialistas–, estos

principios ponen especial énfasis en mejorar la capacidad de las máquinas para encontrar y utilizar los datos automáticamente, además de respaldar su reutilización por parte de los individuos.

Se desarrollan bajo estándares que conectan a investigadores, editores y repositorios de datos para habilitar los datos (localizables, accesibles, interoperables y reutilizables) a gran escala. Apunta también a acelerar el descubrimiento científico y mejorar la integridad, transparencia y su reproducción.

## **El marcaje de datos y la literatura**

Para llevar a cabo el desarrollo de un Repositorio Institucional de datos, o en su caso literatura, deben discutirse una serie de normativas, técnicas y elementos –como son los metadatos que definirán el registro y la especificidad de cada uno de los objetos digitales albergados en el RI–. Estos son indispensables para que los contenidos sean recuperables mediante herramientas de búsqueda en Internet.

A través de los metadatos se crean datos estructurados que describen el contenido, calidad y condición y otras características de un dato y además expliquen y localicen un recurso de información, y, una vez recuperado, usarlo y administrarlo. Al trabajar con la estructura técnica del repositorio se debe hacer mediante estándares internacionales con el fin de poder establecer una comunicación entre las diversas plataformas con las decidamos interactuar.

Es importante dejar claro que se pueden emplear diferentes tipos de metadatos según la naturaleza del Repositorio ya sea de Datos o Literatura. En los cuales se deben tomar en cuenta la descripción de contenido, calidad, condiciones, entre otros, así como normativas, técnicas y elementos que definirán el registro y la especificidad de cada uno de los objetos digitales albergados. Estos son indispensables para que los contenidos sean recuperables mediante herramientas de búsqueda en Internet.

La tipificación de los elementos en los metadatos, generalmente se reconocen de acuerdo a la descripción del objeto digital: administrativos, descriptivos, de preservación, técnicos y de uso. Debemos asegurarnos que los metadatos mantendrán su capacidad de interpretación y utilidad para que los registros puedan almacenarse y conservarse a lo largo del tiempo.

Las funciones del marcaje facilitan la organización de los recursos basados en la audiencia permitiendo organizarlos mediante una identificación digital (URL, PURL o DOI) y estructura, asegurando su preservación. Y pueden ser expresados en XML, lenguaje independiente del software e intercambiable entre distintos sistemas. Se expresa a partir de una Definición de Tipos de Documentos (DTD), especificando las etiquetas que se emplearán.

La aplicación de estos es en objetos digitales como videos, imágenes, mapas, recursos educativos, además de recursos geoespaciales, datos geográficos, entre otros.

Por lo que es necesario una interconexión para mantener una sincronización de la información, accesible para el usuario mediante los llamados *web services* que manejan servicios como:

- Interoperabilidad entre aplicaciones de software, independientemente de sus propiedades o de las plataformas sobre las que se instalen.
- Fomentan los estándares y protocolos basados en texto, que hacen más fácil acceder a su contenido y entender su funcionamiento.
- Apoyan los servicios eeb (HTTP).
- Permiten que los servicios y el software de diferentes compañías se combinen fácilmente para proveer servicios necesarios para el tratamiento de los datos.

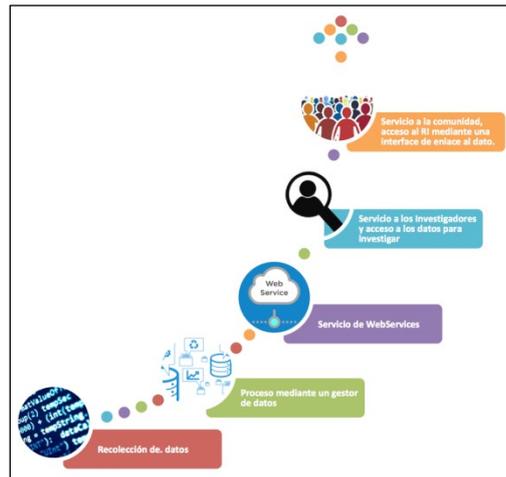


FIGURA 1. Estructura del software con el dato. Elaboración propia

Para las Ciencias de la Tierra los datos tienen condicionantes, por mencionar algunas de ellas, es que se componen de varios archivos en diversos formatos, por lo que es necesario relacionar el dato que está siendo registrado, en ocasiones por nacimiento o naturaleza se relaciona con otro recurso. Es decir, se tiene que llevar a cabo una asociación. Por lo que además se debe considerar las siguientes características:

- Descripción de los datos y formato y versión.
- Cómo se adquieren y procesan.
- Relación con otros datos.
- Metadatos de relación o con los que se requieran para que los datos posean sentido y guarden su contexto.
- Políticas para compartir, acceder y reusar los datos.
- Citación.
- Responsabilidad y roles.
- Derechos de autor.
- Almacenamiento.
- Licencia abierta para acceder libremente a la información.
- Los datos deberán estar estructurados en un formato legible por máquina.

- Se debe depositar en formatos no propietarios, es decir, datos abiertos, estructurados y sin propietarios y con nombres.
- Emplear URI con otras fuentes de datos y aplicaciones, facilitando la identificación.

En cuanto al marcaje empleamos metadatos relacionados con las Ciencias de la Tierra y además los metadatos para datos para los repositorios institucionales ver. 3.0.80: Lineamientos Específicos para Repositorios del Apéndice 2.<sup>9</sup>

Para la parte literatura fue el modelo de Dublín Core, apegado a las estipulaciones de los Lineamientos Específicos para repositorios<sup>10</sup>, publicados en el sitio del Repositorio Nacional, haciendo uso de la notación de Dublín Core básica, no incluyendo calificadores así como también en las Directrices de OpenAIRE 3.0.

Con sus elementos opcionales y repetibles, codificado en HTML, XML y RDF, formatos más divulgados y aceptados para representación y recuperación de información en la web, como se presentan en el siguiente cuadro:

---

<sup>9</sup> México. Infotec. Interoperabilidad con el Metabuscaor del Repositorio Nacional Versión 3.1., publicado el 05-Oct-2018.

<sup>10</sup> Op. cit.



Considerar además que las versiones anteriores de las directrices, presentan los cambios como un perfil y esquema de aplicación basados en Dublín Core y DataCite o con un nuevo OAI-metadataPrefix, apoyándose en los esquemas de identificación para autores, organizaciones, financiadores, recursos académicos introducción de vocabularios controlados por COAR con el Cumplimiento de la Política de Adquisición de Contenido de OpenAIRE.<sup>11</sup>

## Conclusiones

Para llevar a cabo esta tarea, es imprescindible analizar estructuras técnicas correspondientes para asignar los metadatos adecuados a los datos y que estos puedan contar con un contexto e información relevante y se puedan fusionar con más fuentes de información, generando nuevo conocimiento en otros ámbitos y a su vez crear valor adicional y, por otro lado, en cuanto a la literatura, ofrecer producción científica en abierto al alcance de la sociedad.

Como se puede ver a lo largo del trabajo es evidente que se debe llevar a cabo un proceso y análisis de los contenidos, además para un correcto marcaje es importante mencionar que dentro de las Ciencias de la Tierra, el Mar y la Atmósfera se genera una gran cantidad de información en lapsos muy cortos, por lo que se necesitan infraestructuras en donde la gestión de los datos como de contenidos sirvan para suministrar información sobre lo que se produzca en datos y en la parte de la literatura, lo cual facilitará su descripción y accesos, asegurando la calidad y el ofrecimiento de las condiciones de uso. Con ello se mostrará el contexto en donde se encuentra la información, garantizando su disponibilidad y preservación.

Es necesario trabajar junto con otras comunidades de partes interesadas (investigadores, universidades y financiadores de la ciencia) con el fin de continuar con este misión de ponerlo en abierto y valorar aun más inclusión,

---

<sup>11</sup> México. Infotec. Interoperabilidad con el Metabuscador del Repositorio Nacional Versión 3.1., publicado el 05-Oct-2018.

diversidad, confianza e innovación. Así mismo, los metadatos deben estar eficientemente conformados y establecidos bajo parámetros internacionales que ofrezcan una garantía de cosecha e intercambio de información y datos.

El llevar cabo una gestión correcta y adecuada en los repositorios, beneficia en gran medida la representación de los proyectos de investigación y ofrecen un dominio específico o un ámbito que se requiere comprender para un proceso de investigación que necesitan los especialistas. Los investigadores no conocen los esquemas de metadatos ni la lógica del acceso, por lo que nuestra contribución como profesionales de la información es primordial dentro de esta actividades tanto académicas como científicas para seguir en las líneas de la ciencia abierta.

## Bibliografía

- AMERICAN GEOPHYSICAL UNION (2019a). Declaration adopted by the American Geophysical Union. <https://ethics.agu.org/>
- AMERICAN GEOPHYSICAL UNION (2019b). Repository Finder. <https://repositoryfinder.datacite.org/>
- ARMENDÁRIZ SÁNCHEZ, S.; SOSA ZARAGOZA, P.; SOSA JIMÉNEZ, M. (2021). Los datos para la investigación en los repositorios institucionales y su impacto científico y social: los RI en Ciencias de la Tierra. Barcelona: Octaedro. Colección: Universidad; pp. 153-164 .
- ATLASSIAN CONFLUENCE OPEN SOURCE PROJECT. <https://wiki.lyrasis.org/display/DSDOC6x/OAI+2.0+Server>
- CASTRO ESCAMILLA, M. (2020). Desarrollo de un Repositorio de datos científicos de apoyo a la investigación: el caso de las ciencias de la Tierra [tesis de maestría]. Universidad Nacional Autónoma de México. <http://132.248.9.195/ptd2020/agosto/0802846/Index.html>
- CASTRO ESCAMILLA, M. (2021) Las declaraciones, principios e iniciativas de la Ciencia Abierta: una visión. En: Ciencia Abierta Opciones y experiencias para México y Latinoamérica. Barcelona: Octaedro. Colección: Universidad; pp. 79-90.

CONACYT (2017). Lineamientos generales de Ciencia Abierta: Conacyt.

CONACYT (2019). El Repositorio Nacional de México.  
<http://www.repositorionacionalcti.mx/>

DATA CITE (2021). DataCite Metadata Schema. <https://schema.datacite.org/>

HODSON, S. (2021). Turning FAIR into Reality Report and Action Plan.  
[https://www.eoscpilot.eu/sites/default/files/18-11-22-hodson-turning\\_fair\\_into\\_reality.pdf](https://www.eoscpilot.eu/sites/default/files/18-11-22-hodson-turning_fair_into_reality.pdf)

LYRISIS. DSpace: resource. <https://duraspace.org/dspace/resources/>

MÉXICO. INFOTEC. Interoperabilidad con el Metabuscador del Repositorio Nacional Versión 3.1., Publicada el 05-Oct-2018.

SCHIRRWAGEN, J. (2018). OpenAIRE Guidelines for institutional and thematic repository managers 4.0 November 30, 2018.

TORRES-SALINAS D., NICOLÁS; ROBINSON GARCÍA; ALVARO CABEZAS-CLAVIJO (2012). Compartir los datos de investigación en ciencia: introducción al datasharing. *El Profesional de la Información*, 2(12).

UNAM. IGEF (2019). Repositorio Institucional de Recursos Geotérmicos para México.  
<http://rigeotermia.geofisica.unam.mx/jspui/>

UNAM. SSN (2019). Repositorio Institucional de Datos del Servicio Sismológico Nacional.  
<http://rigeofisica.ssn.unam.mx/jspui/>

# Planejamento e incubação de repositórios de dados em ICTs brasileiras

## **CATERINA GROPOSO PAVÃO**

Universidade Federal do Rio Grande do Sul (UFRGS) (Brasil)

[caterina@ufrgs.br](mailto:caterina@ufrgs.br)

## **CAROLINA HOWARD FELICISSIMO**

RNP

[carolina.felicissimo@rnp.br](mailto:carolina.felicissimo@rnp.br)

## **RAFAEL PORT DA ROCHA**

Universidade Federal do Rio Grande do Sul (UFRGS) (Brasil)

[rafael.rocha@ufrgs.br](mailto:rafael.rocha@ufrgs.br)

## **RENE FAUSTINO GABRIEL JUNIOR**

Universidade Federal do Rio Grande do Sul (UFRGS) (Brasil)

[rene.gabriel@ufrgs.br](mailto:rene.gabriel@ufrgs.br)

## **SAMILE ANDREA DE SOUZA VANZ**

Universidade Federal do Rio Grande do Sul (UFRGS) (Brasil)

[samile.vanz@ufrgs.br](mailto:samile.vanz@ufrgs.br)

## **SÔNIA ELISA CAREGNATO**

Universidade Federal do Rio Grande do Sul (UFRGS) (Brasil)

[sonia.caregnato@ufrgs.br](mailto:sonia.caregnato@ufrgs.br)

## **RESUMO**

O texto relata as atividades do GT Rede de Dados de Pesquisa (GT-RDP) no âmbito do Edital de 1/06/2021, lançado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), o

Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a Rede Nacional de Ensino e Pesquisa (RNP). O edital abrange a implementação de um repositório de dados com por quatro instituições, Centro Brasileiro de pesquisas Físicas (CBPF), Universidade Federal do Ceará (UFC), Universidade Federal de Goiás (UFG) e Universidade Federal Rural do Amazonas (UFRA). O objetivo geral da equipe do GT-RDP, responsável pela execução do projeto, foi atender as quatro Instituições de Ciência, Tecnologia e Inovação (ICTs) nas atividades desenvolvidas nas vertentes tecnológica e conceitual, assim como desenvolver atividades de capacitação. A prática de atendimento das quatro ICTs aconteceu por meio de reuniões semanais, onde foram realizadas orientações e direcionamento para a implementação dos repositórios. Utilizou-se o software Dataverse, com base no modelo OAIS e princípios FAIR. O momento de finalização das atividades é dedicado à definição de políticas para funcionamento do repositório. O projeto encontra-se em fase final dos testes e em breve passando para a fase de produção e disponibilização pública.

#### **PALABRAS CLAVE**

Ciência aberta; Dados de Pesquisa; Repositórios de dados de pesquisa.

*Open Science; Research Data; Research Data Repositories.*

## **Introdução**

Com base nos princípios essenciais de liberdade acadêmica, integridade nas pesquisas e excelência científica, a ciência aberta estabelece um novo paradigma que integra práticas científicas de reprodutibilidade, transparência, intercâmbio e colaboração resultante da maior abertura de conteúdos científicos, ferramentas e processos, com o objetivo de tornar o conhecimento científico disponível, acessível e reutilizável, para aumentar a colaboração

científica e o compartilhamento de informações e para abrir os processos de criação, avaliação e comunicação do conhecimento científico (UNESCO, 2021). Os dados de pesquisa abertos são um dos pilares da ciência aberta, usados como fontes primárias para apoiar as pesquisas e a produção do conhecimento, são utilizados como evidências no processo de pesquisa e são comumente aceitos na comunidade científica como necessários para validar as descobertas e resultados da pesquisa (RESEARCH, 2022).

A necessidade de gestão dos dados de pesquisa contribuiu para a criação dos repositórios de dados que, na sua maioria, estão vinculados às universidades e instituições de pesquisa e contribuem para assegurar que os dados sejam publicados e disponibilizados para a comunidade científica com o mínimo de restrições possíveis. Os repositórios de dados de pesquisa são ambientes digitais implementados com infraestrutura para dar suporte aos pesquisadores na gestão e na disponibilização de dados o que potencializa a reutilização por outros pesquisadores.

As universidades internacionais têm implementando repositórios de dados para armazenar e disponibilizar os dados de pesquisa da sua comunidade. As universidades brasileiras, alinhadas com essa tendência, seguem nessa direção. Conforme levantamento realizado por GABRIEL JUNIOR et al. (2019), foram identificados no Brasil 15 repositórios de dados de pesquisa, sendo quatro de abrangência internacional, ou seja, envolvendo várias instituições, com pelo menos uma do Brasil, e 11 de abrangência nacional, sendo que dois desses são do tipo multi-institucional. As temáticas atendidas por esses repositórios são voltadas às Ciências Sociais Aplicadas, às Ciências Agrárias, às Ciências Exatas e da Terra, às Ciências Biológicas e à Engenharia. Atualmente no RE3data<sup>1</sup> estão registrados 14 repositórios brasileiros.

---

<sup>1</sup> [https://www.re3data.org/search?query=&countries\[\]=BRA](https://www.re3data.org/search?query=&countries[]=BRA)

Baseados nesse contexto o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a Rede Nacional de Ensino e Pesquisa (RNP), por meio do Edital de 1/06/2021, convidaram Instituições de Ciência, Tecnologia e Inovação (ICTs) a colaborar para o fomento à Ciência, Tecnologia e Inovação por meio da produção, socialização e integração do conhecimento científico e tecnológico. A ideia foi realizar a colaboração por meio do planejamento e implementação de ações de pesquisa e desenvolvimento (P&D) e transferência de conhecimentos relacionados à criação de repositórios de dados de pesquisa. A equipe de execução do projeto foi composta pelos pesquisadores do GT-RDP.

O GT-RDP é responsável pelo projeto Rede de Dados de Pesquisa Brasileira (RDP Brasil), que vem sendo desenvolvido por pesquisadores ligados ao Centro de Documentação e Acervo Digital da Pesquisa, unidade da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul (CEDAP/FABICO/UFRGS). Em 2017 o grupo iniciou suas atividades, com o objetivo inicial de realizar um levantamento de informações que subsidiassem as primeiras ações no âmbito do Acesso Aberto da Dados de Pesquisa (AADP) no país, visando à identificação de práticas de AADP e o desenvolvimento de um futuro serviço de acesso aberto. Até 2019 o GT- RDP contou com a contribuição do Centro de Ciências Computacionais da Universidade do Rio Grande (C3/FURG).

A execução dos objetivos iniciais do projeto permitiu reunir informações acerca do contexto brasileiro. O grupo desenvolveu uma survey com 4.676 respondentes (CAREGNATO *et al.*, 2019), revelando que, a pesar do grande interesse pelo tema, evidenciado pela prevalência de variáveis relacionadas ao compartilhamento e ao uso de dados e aos repositórios institucionais, não há clareza por parte dos pesquisadores sobre os principais tópicos relacionados. Em média, 49,36% dos respondentes nunca utilizou dados compartilhados por outros grupos; e quase um quarto dos respondentes (23,49 %) nunca compartilhou dados de pesquisa, com percentuais que variam de 14,18 % na

Linguística, Letras e Artes até 35,29 % nas Ciências da Saúde. Entende-se que a gestão e a disponibilização de dados de pesquisa é a chave para que esses possam ser utilizados por pesquisadores brasileiros ou estrangeiros em futuras pesquisas. Apesar de 57,79 % dos respondentes armazenarem parte dos seus dados em arquivos em papel e 66,80 % guardarem todos os dados em seu computador pessoal, há um entendimento geral de que os dados devam ser compartilhados. Fica evidente o problema de infraestrutura para o armazenamento e o compartilhamento dos dados de pesquisa nas instituições brasileiras ao analisar que 71,31 % dos respondentes apontaram que a instituição de trabalho nunca oferece orientações sobre o tema e que 78,56 % deles responderam também que a instituição não oferece pessoal especializado para preparar os dados a serem compartilhados.

Munidos de diversas informações, o grupo passou a dedicar-se ao estudo de estratégias de planejamento para implantação de repositórios, preservação digital e certificação. Este trabalho relata as atividades de planejamento e incubação de repositórios em quatro ICTs atendidas pelo Edital de 1/06/2021, conforme detalhado a seguir.

## O projeto

Considerando a crescente demanda de compartilhamento de dados de pesquisa motivada pelo novo paradigma da Ciência Aberta, as diretrizes produzidas pelo Grupo de Trabalho formado para discutir uma proposta de Política Nacional para Ciência Aberta, as recentes recomendações da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) sobre acesso a dados de pesquisas científicas com financiamento público e o trabalho realizado pelo IBICT e RNP no âmbito do Compromisso três do 4º Plano de Ação Nacional para Governo Aberto (OGP), o Plano de Trabalho executado teve como objetivo principal o apoio à implantação de repositórios de dados de pesquisa nas quatro ICTs selecionadas pelo Edital. Como compromisso, cada partícipe deveria aplicar recursos humanos e financeiros,

computacionais, assinaturas ou licenças de software, dentre outros, para a execução do plano de trabalho.

Para pôr em prática o plano de trabalho as ICTs se comprometeram a implementar, até o final do projeto piloto, um repositório de dados com, pelo menos, três conjuntos de dados de pesquisadores que estivessem dispostos a colaborar com a iniciativa. As instituições selecionadas e a área dos datasets foram:

- 1. Centro Brasileiro de pesquisas Físicas (CBPF)**, o piloto do repositório de dados de pesquisa incubado conterá datasets, previamente selecionados, dos laboratórios: a) Laboratório Multiusuário de Nanociência e Nanotecnologia – LABNANO; b) Laboratório de Ressonância Magnética Nuclear – RMN; c) Laboratório de Raios X - RX e Nanomagnetismo e d) Espectroscopia Mössbauer. Laboratório de Magnetismo Aplicado. A equipe é formada por um físico e pesquisador do Centro, juntamente com uma bibliotecária e um responsável pelo setor de TI.
- 2. Universidade Federal do Ceará (UFC)**, o piloto do repositório de dados conterá datasets das seguintes áreas: a) Engenharia de Produção, Inteligência Computacional e Otimização; b) Fisioterapia, Tecnologias assistivas, Bioengenharia, Engenharia Biomédica; c) Letras, Estudos da Língua Inglesa, Literatura, Tradução, Fonética e Multilinguismo e d) Finanças, Economia, Engenharia de Transportes. A equipe é formada por uma bibliotecária.
- 3. Universidade Federal de Goiás (UFG)**, o piloto do repositório de dados de pesquisa conterá datasets, das seguintes áreas: a) Ciências Biológicas-Botânica / Ciências Agrárias-Fitotecnia-Tecnologia de Sementes / Ciências Agrárias-Engenharia Florestal e Recursos Florestais-Silvicultura; b) Ciências da Saúde, Cardiologia; c) Ciências Sociais, Sociologia e d) Ciências Humanas, Antropologia, Arqueologia e Museologia. A equipe é formada por professores do Departamento de Ciência da Informação,

gerentes de TI, e professores vinculados a administração central da Universidade.

- 4. Universidade Federal Rural do Amazonas (UFRA)** o piloto do repositório de dados de pesquisa conterà datasets das seguintes áreas: a) Fisiologia Animal - Reprodução de Primata na Amazônia legal Fisiologia Reprodutiva de Animais Selvagens e Domésticos na Amazônia; b) Recuperação de Áreas Degradadas, Manejo florestal em concessões federais no Pará, Ecologia, manejo e produção florestal das várzeas do Marajó, Avaliação de componentes químicos de diferentes espécies florestais armazenadas, passíveis de contaminação no solo e c) Reprodução animal- Desenvolvimento de um novo diluente crioprotetor a base de castanha-do-pará /brasil, Suplementação para reduzir a idade a puberdade de novilhas zebuínas em pastagens tropicais. A equipe é formada por uma bibliotecária, por um responsável pela TI e pela superintendente a rede da instituição.

## Execução do projeto

O objetivo geral da equipe do GT-RDP, responsável pela execução do projeto, foi atender as quatro ICTs nas atividades desenvolvidas nas vertentes tecnológica e conceitual, assim como desenvolver atividades de capacitação. Mais especificamente, proporcionar suporte tecnológico e conceitual para execução das atividades; capacitar as ICTs na vertente conceitual e tecnológica e participar em atividades de disseminação dos resultados para criação de repositórios de dados. Ainda, como objetivos secundários, apoiar instituições de ensino e pesquisa brasileiras na abertura de dados de pesquisa; fomentar o debate acerca da importância do desenvolvimento de repositórios de dados de pesquisa em instituições brasileiras; apoiar atividades de P&D e experimentação em repositório de dados de pesquisa a partir de estudos de caso com datasets de diferentes áreas do conhecimento; apoiar ações da RNP no processo de descoberta e estruturação de serviços para a implantação de

repositórios de dados de pesquisa; apoiar ações nacionais voltadas para a elaboração de uma política para ciência aberta; ampliar as instituições participantes do Consórcio Nacional para Ciência Aberta (CoNCienciA) com a inclusão das ICTs selecionadas, consórcio que fornecerá identificadores persistentes associados a seus conjuntos de dados (DOI); e finalmente, planejar a estruturação de uma Rede Federada Nacional de Repositórios de Dados de Pesquisa, com a participação dos repositórios de dados do IBICT, CNPq, RNP e de um conjunto de instituições de pesquisa convidadas, entre elas as ICTs selecionadas por meio do edital.

Para alcançar esses objetivos foi necessário capacitar as equipes das ICTs selecionadas em diversos aspectos relacionados à gestão de repositórios de dados de pesquisa, sobre a instalação, manutenção, operação, utilização e governança de repositórios de dados de pesquisa. Relatam-se a seguir alguns dos princípios que guiaram esse processo. A implantação de um repositório de dados é orientada por uma série de modelos de referência e princípios, como o Modelo de Referência para Repositórios, Open Archival Information System (OAIS) e os princípios de compartilhamento FAIR e de citação. Além disso, para o desenvolvimento de um repositório, é importante considerar critérios para construção de repositórios digitais confiáveis.

OAIS é um modelo de referência que proporciona uma linguagem comum para facilitar a discussão entre comunidades interessadas no desenvolvimento de repositórios com responsabilidade em preservação digital (Consultive..., 2012). É composto pelas especificações do ambiente do repositório, do modelo de dados das informações a serem preservadas e do modelo funcional do repositório. Um Repositório confiável é aquele cuja missão é fornecer acesso de longo prazo a recursos digitais gerenciados. Um repositório com essas características aceita a responsabilidade pela manutenção a longo prazo dos recursos digitais, projeta seus sistemas de acordo com as convenções e os padrões comumente aceitos, estabelece metodologias para avaliação de sistemas que atendem às expectativas de confiabilidade da comunidade, cumpre suas responsabilidades com depositantes e usuários de forma aberta e

explícita, e permite que sejam auditadas e medidas suas políticas, práticas e desempenho. Critérios para verificação e certificação de repositórios digitais confiáveis têm como base e medem a proximidade do repositório com o Modelo de Referência OAIS. Os principais instrumentos para certificação de repositórios digitais são Core Trusted Seal (Becker,2009) e ISO 16363 27.

A garantia de armazenamento e acessibilidade dos dados a longo prazo é uma exigência importante e nesse sentido, adoção dos princípios TRUST (Transparency, Responsibility, User focus, Sustainability and Technology) foi trabalhada com as quatro ICTs atendidas pelo projeto, conforme mostra o quadro 1:

<b>PRINCÍPIO</b>	<b>ORIENTAÇÃO PARA OS REPOSITÓRIOS</b>
T (transparência)	Ser transparente sobre os serviços de repositório específicos e acervos de dados
R (responsabilidade)	Ser responsável por garantir a autenticidade e integridade dos acervos de dados e pela confiabilidade e persistência de seu serviço.
U (foco no usuário)	Para garantir que as normas de gerenciamento de dados e expectativas das comunidades de usuários sejam atendidas.
S (sustentabilidade)	Para sustentar serviços e preservar acervos de dados para o longo prazo.
T (tecnologia)	Para fornecer infraestrutura e recursos para suporte a serviços seguros, persistentes e confiáveis.

QUADRO 1. Princípios TRUST para repositórios digitais  
Fonte: Lin et al. (2020)

Os Princípios TRUST se constituem como um meio para agilizar a comunicação com todas as partes interessadas e servem como um guia para que os repositórios demonstrem transparência, responsabilidade, foco no usuário, sustentabilidade e tecnologia (LIN et al., 2020). Aliados a eles, os princípios FAIR e de citação orientam no sentido de promover o compartilhamento dos dados (FORCE11, 2017). Os princípios FAIR indicam que

dados devem ser localizáveis (Findable), acessíveis (Accessible), interoperáveis (Interoperable) e reusáveis (Reusable). Para dados serem FAIR, necessitam identificadores únicos, persistentes e globais (F). Devem ser descritos por metadados indexáveis e ricos (F), representados em linguagens formais (I), aceitos pela comunidade (R), com atributos relevantes, precisos e úteis ao contexto (R), incluindo metadados de proveniência (R) e usando vocabulários controlados que seguem princípios FAIR (I). Estes dados devem ser recuperáveis pelo seu identificador através de um protocolo de comunicação padronizado, aberto e gratuito (A). Também devem ser acompanhados de licenças claras e acessíveis (R), e referências qualificadas devem ligar (meta)dados para enriquecer o conhecimento sobre os mesmos (I).

Identificadores persistentes viabilizam a referência de longa duração aos objetos digitais, à medida que permitem que estes objetos sejam identificados de forma independente de suas localizações. O Digital Object Identifier (DOI) e Handle System são os principais serviços de identificadores persistentes usados em repositórios digitais. DOI é um sistema que agrega metadados ao Handle System, permitindo a adição de semântica aos identificadores. Especifica um conjunto nuclear de metadados, que podem ser expandidos com novos elementos para atender domínios específicos. O DOI possui um modelo de negócio em que o registro é realizado por meio de uma federação de agências. Essas agências atendem a domínios específicos, e usam esquemas de metadados específicos para atender às necessidades desses domínios.

O software Dataverse foi utilizado como solução tecnológica para repositório de dados de pesquisa em uma prova de conceito de instanciação do modelo OAIS para o objeto de estudo proposto. A escolha do software Dataverse baseia-se, principalmente: no estudo comparativo de soluções tecnológica para compartilhamento de dados realizado pelo GT-RDP (Rocha, 2018); na aceitação e direcionamento do IBICT pelo uso do software Dataverse; e, pela experiência adquirida pelas equipes da RNP e IBICT na instalação e operação do Dataverse, que será usada para o suporte das atividades previstas do projeto.

De acordo com ROCHA (2018), o Dataverse possui recursos para configuração de vários tipos de ambientes de repositório, incluindo hierarquias organizacionais e políticas de gestão distintas para unidades ou grupos, incluindo esquemas de metadados e licenças e por isso apresentou-se como a melhor opção entre os softwares analisados no estudo para a implementação de repositórios de dados de pesquisa.

Alguns esquemas de metadados já vêm pré-definidos na instalação no Dataverse, compostos por metadados de citação e outras variáveis. O Dataverse disponibiliza os esquemas para as áreas de: Ciências Sociais e Humanidades, Geoespaciais, Astronomia e Astrofísica, Ciências da Vida, Revistas (Journals). Esses esquemas foram criados pensando no Data Documentation Initiative (DDI) e DataCite. O Esquema de Citação é obrigatório, embora nem todos os metadados dentro desse esquema sejam obrigatórios. É possível manter um conjunto mínimo mandatório do esquema e citação, composto pelo Identificador, Criador (Autor), Título, Publicador, Ano de publicação e tipo de recurso. É possível mesclar o Esquema de Citação com os específicos das áreas do conhecimento previstas, criando um perfil de aplicação. Por exemplo: Esquema de Citação + Esquema de Ciências Sociais e Humanidades. Como os campos do Esquema de Citação são pré-definidos, eles já possuem mapeamento com o DDI (crosswalks). Os metadados do Esquema de Citação no Dataverse incluem aqueles que descrevem os dados mais gerais de um documento. Dentre vários outros possíveis, elencou-se a seguir 25 elementos para exemplificar:

<b>METADADOS NO DATAVERSE</b>	<b>DESCRIÇÃO EM PORTUGUÊS</b>
title	Título
otherId	Outra identificação
author	Autor
datasetContact	Contato do conjunto de dados
dsDescription	Descrição
subject	Assunto
keyword	Palavra-chave
topicClassification	Classificação de tópicos
publication	Publicação
notesText	Notas de texto
Language	Idioma
producer	Produtor
contributor	Contribuidor
grantNumber	Número de concessão
distributorName	Nome do distribuidor
depositor	Depositante
dateOfDeposit	Data do depósito
timePeriodCovered	Cobertura temporal
dateOfCollection	Data de coleta
kindOfData	Tipo de dados
series	Série
software	Programa
relatedMaterial	Material relacionado
dataSources	Fontes de dados
accessToSources	Acesso a fontes

QUADRO 2. Esquema de citação do Dataverse. Fonte: Criação do GT-RDP

A capacitação das quatro ICTs é realizada por meio de reuniões semanais, onde acontece a apresentação de boas práticas e fundamentos teóricos, além da discussão das atividades práticas e de experimentação desenvolvidas pelo grupo. Os atendimentos seguem uma vertente conceitual, onde foram discutidos, entre outros temas, a estrutura de organização do repositório, os metadados de descrição, definição de papéis e grupos, fluxos de submissão, políticas para o repositório, identificadores persistentes e requisitos de certificação.

As questões técnicas de instalação e configuração do software Dataverse foram contempladas em atendimentos da vertente tecnológica. Ao longo deste processo, desenvolveu-se um checklist dos itens básicos a serem operacionalizados na instanciação de um repositório:

<input type="checkbox"/>	Ambiente de instalação (min. 4 núcleos, 8Giga RAM, 1Tera disco)
<input type="checkbox"/>	Pré-requisitos (banco de dados PostGres, Java 11, plataforma de pesquisa SOLR, Servidor de aplicativo PAYARA, JQ para Json, ImageMagick para imagens e RServer para dados estatísticos)
<input type="checkbox"/>	Definição da URL de acesso (http)
<input type="checkbox"/>	Serviço de monitoramento (Ex: Google Analytics)
<input type="checkbox"/>	Instalação de uma chave de segurança
<input type="checkbox"/>	Servidor para envio de e-mail (SMTP) – podendo ser servidor externo como o Google
<input type="checkbox"/>	Serviço de Identificador persistente (DOI do DataCite ou Handle), deve-se definir se a atribuição do DOI será com conjunto de Datasets ou por Arquivos (gera maior número de DOIs)
<input type="checkbox"/>	Definição da forma de autenticação (bultin: do próprio Dataverse ou Shibboleth – federada)
<input type="checkbox"/>	Rotinas de backup e restauração do Dataverse (Postgres e Payara – arquivos)
<input type="checkbox"/>	Customizações

<input type="checkbox"/>	Customização da Página inicial (opcional)
<input type="checkbox"/>	Página de políticas
<input type="checkbox"/>	Página do Guia para usuários
<input type="checkbox"/>	Logo da instância (altura recomendada até 45pixel)
<input type="checkbox"/>	Criação da página sobre
<input type="checkbox"/>	Implementação dos idiomas de navegação
<input type="checkbox"/>	Customizações (adicionais)
<input type="checkbox"/>	Perfil de aplicação – perfil de metadados personalizado ( <i>Schema</i> )
<input type="checkbox"/>	Configurações para armazenamento elástico em nuvem (AWS, Azure, Google)

QUADRO 3. Checklist para implementação tecnológica

Fonte: Criação do GT-RDP

Nas configurações do Dataverse ressalta-se que o registro de um Dataset pode ocorrer de duas maneiras, uma com a identificação de cada arquivo e sua versão, o que na prática gera uma quantidade grande de registros, por consequência aumentando o seu custo<sup>2</sup>, ou seja, ao registrar um Dataset com 1.000 arquivos, serão gerados 1.001 DOIs diferentes, um para cada arquivo e outro para o Dataset. A outra forma é atribuição para o Dataset, tendo os arquivos herdados o DOI do Dataset, gerando apenas um DOI, menor custo. Essa opção deve ser avaliada pelos gestores na implementação do repositório.

A estratégia de instanciamento do Dataverse foi o uso de dois servidores, uma para os testes denominado de desenvolvimento e outro de produção com acesso público, no servidor de desenvolvimento foi utilizado do DOI FAKE<sup>3</sup>. Observou-se que a se atender os requisitos e a obtenção de das informações previamente, a instalação pelas ICTs ocorreu em três semanas. A maior

---

<sup>2</sup> A cobrança do DOI pelo DataCite é realizada previamente com a contratação de pacotes anuais, com valor próximo de 0,80 euros por registro de DOI. <https://datacite.org/feemodel.html>

<sup>3</sup> DOI registrado no ambiente de testes do DATACITE

dificuldade foi na definição das pessoas que compõem a equipe tecnológica, pois no início houve uma grande rotatividade. Outro ponto a destacar está na documentação dos problemas enfrentados na instalação e principalmente nos registros de senhas que deve estar disponível para os gestores do projeto.

Ressalta-se também a vertente tecnológica depende muito da vertente teórica para as definições da instalação e customização, e que reuniões em conjunto são essenciais para o bom desenvolvimento dos repositórios, com as definições das políticas, pois o Dataverse permite muitas configurações. Destaca-se também que as customizações realizadas são realizadas pelas APIs e BASH, sem a necessidade de alteração do código fonte, ou seja, em novas atualizações todas as configurações permanecem compatíveis.

Além dos atendimentos realizados nas reuniões semanais, a produção de vídeos mostrou-se uma aliada para disseminar conhecimentos, tanto para atendimento às ICTs como para fomentar o debate e o entendimento a acerca da importância e do desenvolvimento de repositórios de dados de pesquisa em instituições brasileiras. O grupo RDP Brasil produziu 7 vídeos relativos aos conceitos básicos de ciência aberta, metadados e operacionalização do Dataverse, disponíveis em

<https://www.youtube.com/playlist?list=PLSk7nwl2wEh9CBEEqFRB6E5BqXiiC7HNP>

## Conclusão

O projeto encontra-se em fase de finalização dos testes e as quatro ICTs já estão com repositórios instalados, operando a admissão (ingest) de datasets e descrição dos mesmos, os próximos passos estão relacionados a implementação dos repositórios em produção. Paralelamente ao desenvolvimento das questões técnicas, o grupo vem trabalhando nas políticas do repositório e fluxos de submissão e uso. Pretende-se, ao longo do segundo semestre, incentivar uma Rede Brasileira de Repositórios a partir das quatro ICTs, fomentando maior colaboração e apoio entre as instituições parceiras e o

lançamento do LattesData pelo CNPq e o Aléia pelo Ibict. As quatro ICTs participantes do projeto automaticamente tem prioridade de entrada no Consórcio Nacional para Ciência Aberta (CoNCienciA), o que possibilitará ter identificadores persistentes DOI associados a seus conjuntos de dados, com o custo inicial da emissão de DOIs patrocinado pelo CNPq.

Da experiência do GT-RDP na implementação e capacitação de gestores dos repositórios não enfrentou dificuldades no instanciamento do Dataverse e em suas definições, como a política do repositório, sendo fruto de muitas discussões. Entretanto o maior problema está na alteração da cultura da instituição com o pesquisador, principalmente em disponibilizar seus dados e compreensão do ciclo de vida dos dados de pesquisa atendendo os princípios FAIR.

## Bibliografia

- BECKER, C., KULOVITS, H., GUTTENBRUNNER, M., STRODL, S., RAUBER, A., HOFMAN, H. (2009). Systematic planning for Digital. Preservation: evaluating potential strategies and building preservation plans. *International Journal on Digital Libraries*, 10 (4), 133-157. DOI:10.1007/s00799-009-0057-1
- CAREGNATO, S. E.; VANZ, S. A.; PAVÃO, C. M. G.; PASSOS, P. C. S. J.; BORGES, E. N.; GABRIEL JUNIOR, R. F.; AZAMBUJA, L. A. B.; ROCHA, R. P. (2019). Práticas e percepções dos pesquisadores brasileiros sobre serviços de acesso aberto a dados de pesquisa. *LIINC em Revista*, 15, 121-141.
- CONSULTIVE COMMITTEE FOR SPACE DATA SYSTEMS (CCSDS) (2012). Reference Model for an Open Archival Information System (OAIS). // CCSDS 650.0-M-2. Magenta Book, Recommended Practice, Issue 2. <https://public.ccsds.org/Pubs/650x0m2.pdf>
- LIN, D. ET AL. The TRUST Principles for digital repositories. *Scientific Data*, 7(144). <https://www.nature.com/articles/s41597-020-0486-7>
- FORCE11 (2017). Guiding principles for findable, accessible, interoperable and re-usable data publishing version b1.0. 2017. <https://force11.org/info/guiding-principles-for-findable-accessible-interoperable-and-re-usable-data-publishing-version-b1-0/>

GABRIEL JUNIOR, R. F.; ROCHA, R. P.; CAREGNATO, S. E.; PAVÃO, C.G.; PASSOS, P. C. S. J.; BORGES, E. N.; VANZ, S. A. S.; AZAMBUJA, L. A. B. (2019). Acesso aberto a dados de pesquisa no Brasil: mapeamento de repositórios, práticas e percepções dos pesquisadores e tecnologias. *Ciência da Informação*, 48, 87-101.

RESEARCH DATA CANADA (2022). Original RDC Glossary. <https://www.rdc-drc.ca/glossary/original-rdc-glossary/>

ROCHA, R. P. DA; ET AL. (2018). Acesso aberto a dados de pesquisa no Brasil: soluções tecnológicas: relatório 2018. Porto Alegre, UFRGS. 75 p. <http://hdl.handle.net/10183/185126>

UNESCO. Conferência General (2021). 41ª Reunión – París. Proyecto de recomendación sobre la ciencia abierta. [https://unesdoc.unesco.org/ark:/48223/pf0000378841\\_spa](https://unesdoc.unesco.org/ark:/48223/pf0000378841_spa)

Eje temático:

Evaluación y métricas alternativas

[Ver bloque de ponencias en You Tube](#)

# HERA - Herramienta para Enriquecimiento de Recursos Académicos

## JUAN FRANCISCO PORTO

Facultad de Informática, Universidad Nacional de La Plata (UNLP) (Argentina)

## ENZO RUCCI

III-LIDI, Facultad de Informática, Universidad Nacional de La Plata (UNLP) (Argentina)

Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[erucci@lidi.info.unlp.edu.ar](mailto:erucci@lidi.info.unlp.edu.ar)

## GONZALO LUJÁN VILLARREAL

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[gonzalo@prebi.unlp.edu.ar](mailto:gonzalo@prebi.unlp.edu.ar)

## RESUMEN

En la actualidad, determinar la calidad y el impacto de un recurso académico (revista o artículo científico) representa un verdadero desafío para un investigador. Al problema inicial de la existencia de múltiples indicadores y métricas de diversa índole para cada uno de ellos, se ha sumado el crecimiento exponencial en la cantidad de recursos para analizar, gracias al desarrollo tecnológico. En esta ponencia se presenta el diseño y desarrollo de una herramienta web que busca dar respuesta a esta problemática. HERA es una herramienta que apunta a simplificar, agilizar y apoyar el proceso de determinar la calidad y el impacto de un recurso académico. Para ello, HERA consulta múltiples fuentes en tiempo real para luego ofrecer información de un recurso como sus metadatos, su pertenencia a índices y bases de

datos, sus métricas de citas y menciones, e información de la publicación donde figura dicho recurso en caso de que corresponda.

#### **PALABRAS CLAVE**

Artículo científico, revista científica, *paper*, *journal*, bibliometría, bases de datos académicas, bases de datos bibliográficas, cienciometría, recuperación de información, evaluación bibliográfica.

Este trabajo presenta un resumen de: PORTO, JUAN FRANCISCO (2021). HERA: Herramienta para Enriquecimiento de Recursos Académicos. Tesis de grado, Facultad de Informática, Universidad Nacional de La Plata (UNLP). <http://hdl.handle.net/10915/129874>

## **La evaluación de la calidad e impacto en el presente**

A la hora de utilizar la información publicada en revistas, congresos, libros, tesis, etc., los lectores deben contar con criterios que les permitan evaluar el grado de veracidad del contenido de cada artículo. La gran disponibilidad de contenido académico y de espacios donde se publica este contenido hace inviable conocer de antemano los procesos que se utilizan para controlar la calidad, asegurar la transparencia, y obtener métricas en cada uno de ellos. Así es como nace la necesidad de establecer criterios genéricos para determinar y diferenciar la calidad de las publicaciones (GONZALEZ-PARDO et al., 2020).

Una característica buscada intensamente y asociada estrechamente con la calidad tanto de una revista (o *journal*) como de un artículo (o *paper*) es el impacto de estos. Mientras que la calidad de una revista se asocia a los procesos que sigue y a los requisitos que impone para que un artículo sea finalmente publicado, el impacto se asocia a la repercusión que la revista o el artículo tiene en la comunidad científica, usualmente medido en número de citas y menciones (REPISO, 2015). La necesidad de determinar la calidad y el

impacto de las publicaciones se traduce en lo que hoy vemos como sistemas de evaluación, indicadores y métricas de revistas o artículos. Si bien no son una herramienta automatizada en su totalidad, ni mucho menos estandarizada, estos elementos nos permitirían tener valores concretos o numéricos, que posiblemente darían respuesta a preguntas como: ¿es este artículo o esta revista de buena calidad? ¿es este artículo o esta revista de alto impacto? ¿se puede predecir el éxito? ¿se puede entender si es de utilidad? Cabe destacar que los valores obtenidos al medir calidad e impacto poseen un carácter subjetivo muy alto por no ser un estándar y por existir diferentes formas de calcularlos y valorarlos, de acuerdo a las distintas áreas temáticas, organismos de medición y evaluación, métodos utilizados, fuentes de datos, y estrategias y tecnologías utilizadas.

En el caso de las revistas, la calidad suele determinarse a partir de su presencia en determinados sistemas de evaluación, conocidos usualmente como bases de datos bibliográficas. Cabe destacar que, en rigor de verdad, esto no se trata de una métrica por no poseer un valor numérico, con lo cual la interpretación del peso de esta información la hace individualmente cada investigador. Cada una de estas bases de datos tiene sus propios propósitos y criterios de inclusión, y existen en la actualidad una multiplicidad de ellas. Por su parte, la calidad de los artículos habitualmente se asocia a la revista en que fueron publicados, por la necesidad de cumplir sus requisitos de publicación.

En adición a las métricas clásicas basadas en citas desde recursos científicos o académicos, han surgido soluciones alternativas, como, por ejemplo, las altmetrics. Dicho simple, este tipo de métricas engloban e incorporan nuevos aspectos para evaluar contenido, y están muy vinculadas al mundo hiperconectado actual: descargas, menciones en redes sociales, citas en blogs científicos e incorporaciones en sistemas de gestión bibliográfica son algunos de los valores que registran a la hora de medir el impacto de una publicación. Su punto fuerte yace en la facilidad de interpretación visual, sus valores numéricos simples y su conveniente portabilidad. Mientras que su

atractivo es indiscutible, no dejan de introducir nuevos valores a ser interpretados por los investigadores.

Más allá del tipo de métricas o del mecanismo utilizado para evaluar calidad e impacto de una publicación, es evidente que los investigadores poseen hoy en día mucha información a disposición, con más que variados métodos de análisis. Surgen muchas preguntas respecto de los tipos de métricas ya mencionadas, y las problemáticas en general son el natural carácter subjetivo de ellas, o de los sistemas de evaluación, su poca interrelación, y la posibilidad existente de encontrar indicadores bajo las mismas nomenclaturas, pero calculados de formas levemente diferentes en las distintas bases de datos, produciendo valores poco similares. En la práctica, la evaluación de calidad por parte de un investigador que analiza estos valores bajo estos términos se vuelve difícil.

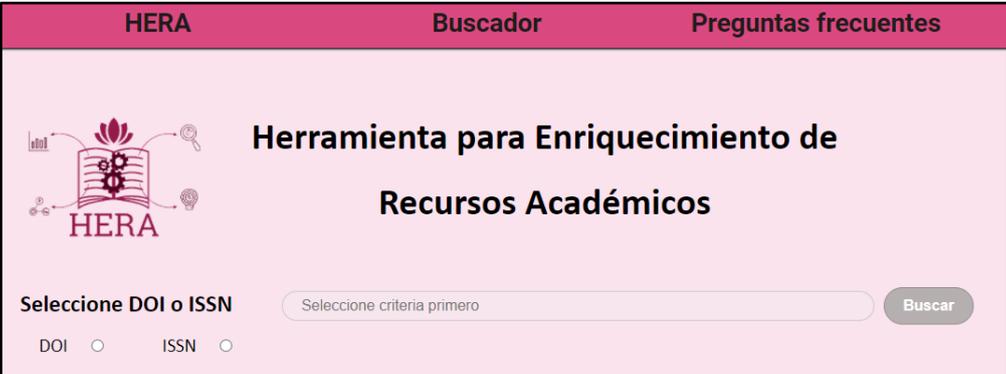
De este análisis de revistas y artículos, sus métricas y los problemas que conlleva analizarlos, entendemos que atacar su carácter subjetivo es un verdadero desafío. El carácter subjetivo no sólo lo otorga las métricas e indicadores, sino que por ser algo natural e inherente a la información, el mismo investigador que la lee y analiza le adjudica un valor personal que podría diferir del original. Existen algunas aplicaciones que intentan dar solución a la problemática descrita, ofreciendo a los investigadores una recopilación de información que permite la evaluación de los recursos académicos con métodos variados. Por ejemplo, el sitio web [¿Dónde lo Publico?](#) recopila información de revistas de ciencias sociales y humanidades en español o portugués. Una opción superadora es la [Matriz de Información para el Análisis de Revistas](#) (MIAR), que da una vista ampliada e integral de métricas de las revistas que posee en su base de datos. Cabe aquí mencionar que MIAR no provee información a nivel de artículo, y Dónde lo Publico provee información a un nivel muy acotado, además de haber sido descontinuada su actualización.

Basándonos en que “a pesar del paso del tiempo, nuestra mejor medida válida de calidad continúa siendo el juicio humano” (LINDSEY, 1989; RHAJEM, 2017), y con el propósito de complementar las propuestas existentes, se

desprende la necesidad de una herramienta automatizada que agilice el trabajo del investigador en su tarea de evaluación individual de contenidos científicos. En lo posible, deberá hacerlo de manera integral mostrando una visión rápida y concisa de las métricas mencionadas, sin que los investigadores deban recorrer decenas de sitios web buscando información sobre publicaciones o trabajos. Bajo esta idea surge la Herramienta para Enriquecimiento de Recursos Académicos, o simplemente HERA, como una solución centralizada que unifica información de revistas y de artículos, automatizando su recolección y brindando un punto de acceso único y simplificado a distintas métricas de calidad e impacto.

## Cómo funciona HERA

HERA es una herramienta web que se encuentra disponible en <http://hera.sedici.unlp.edu.ar>. Su funcionamiento es relativamente sencillo: el usuario cuenta con una barra de búsqueda donde deberá introducir un identificador, que puede ser el ISSN si desea analizar una publicación, o un DOI si desea analizar un artículo (ver FIGURA 1). Luego, deberá hacer clic en el botón “Buscar” que indica a HERA que debe iniciar la búsqueda de información del recurso asociado a dicho identificador.



HERA Buscador Preguntas frecuentes

**Herramienta para Enriquecimiento de Recursos Académicos**

Seleccione DOI o ISSN

DOI  ISSN

FIGURA 1. Interfaz de búsqueda de la herramienta con input de búsqueda y selectores de tipo de identificador (DOI e ISSN)

Una vez iniciada la búsqueda, HERA se conectará de manera concurrente con un conjunto de servicios en línea y sitios web (bases de datos de aquí en adelante), desde los cuales intentará obtener información del recurso de interés. Asimismo, HERA decidirá cuál es la mejor estrategia para obtener información a partir de cada base de datos (por ejemplo, una API REST o *web scrapping*), así como también qué tipo de información puede obtener a partir de cada una de ellos: pertenencia o no a la base de datos, citas, gráficos, altmetrics, etc. La siguiente tabla lista las bases de datos que consulta HERA, así como también las métricas que se obtienen en cada caso:

<b>NOMBRE DE BASE DE DATOS</b>	<b>MÉTODOS DE SELECCIÓN DE CONTENIDO</b>	<b>(PRINCIPALES) MÉTRICAS E INDICADORES PROVISTOS</b>	<b>RECURSO<sup>1</sup> CONSIDERADO</b>
<b>SCOPUS</b>	Proceso de evaluación por junta asesora	Número de citas Predicción de citas Factores de impacto	<i>Journal</i>
<b>WEB OF SCIENCE</b>	Proceso de evaluación por editores expertos	Índice de WoS al que pertenece Información bibliográfica	<i>Journal</i>
<b>DOAJ</b>	Proceso de evaluación de acceso abierto y calidad	Licencias Sello de DOAJ	<i>Paper &amp; Journal</i>
<b>LATINDEX</b>	Proceso de evaluación de calidad (sólo iberoamérica o contenido de la región)	Información bibliográfica Presencia en el catálogo	<i>Journal</i>
<b>RED IBEROAMERICANA</b>	Proceso de evaluación de calidad (sólo Iberoamérica o contenido de la región)	Posición en ranking Información	<i>Journal</i>

---

<sup>1</sup> Si bien se usó el término *Journal*, también pueden estar considerados Conferencias y Libros

NOMBRE DE BASE DE DATOS	MÉTODOS DE SELECCIÓN DE CONTENIDO	(PRINCIPALES) MÉTRICAS E INDICADORES PROVISTOS	RECURSO <sup>1</sup> CONSIDERADO
		bibliográfica Indicadores de calidad editorial y acreditaciones	
<b>CROSSREF</b>	Metadatos estandarizados por miembros de la plataforma	Número de citas Estadísticas de papers por año y autores con ORCID Información bibliográfica	<i>Paper &amp; Journal</i>
<b>SCIMAGO JOURNAL &amp; COUNTRY RANK</b>	Recursos contenidos en Scopus	Posición en ranking global Cuartil H-index del journal Información bibliográfica	<i>Journal</i>
<b>GOOGLE SCHOLAR</b>	Indexación por Google mediante algoritmos de <i>machine learning</i>	Número de citas H-index de autores	<i>Paper</i>
<b>MICROSOFT ACADEMIC</b>	Indexación por Bing mediante algoritmos de <i>machine learning</i>	Número de citas Predicción de citas Información bibliográfica	<i>Paper</i>
<b>SEMANTIC SCHOLAR</b>	Indexación mediante algoritmos de <i>machine learning</i> ejecutados sobre bases de datos afiliadas	Número de citas Frecuencia de citas Número de citas influyentes	<i>Paper</i>
<b>ALTMETRIC</b>	Realiza cálculos reconociendo menciones en redes sociales, foros, etc. mediante IA	Puntaje de atención basado en menciones con peso dependiendo el sitio	<i>Paper</i>

NOMBRE DE BASE DE DATOS	MÉTODOS DE SELECCIÓN DE CONTENIDO	(PRINCIPALES) MÉTRICAS E INDICADORES PROVISTOS	RECURSO <sup>1</sup> CONSIDERADO
		Presencia de menciones en redes sociales	
<b>DIMENSIONS</b>	Recopilación de contenidos de bases abiertas (Crossref y PubMed)	Número de citas Número de citas recientes (2 años) Proporción de citas por campo y relativa	<i>Paper</i>

TABLA 1. Resumen de información provista por las bases de datos

## Diseño y desarrollo

Para la implementación de la herramienta se desarrolló una aplicación web, que está dividida en una aplicación front-end, desarrollada con ReactJS, y en una aplicación back-end, desarrollada con NodeJS, como se muestra en la FIGURA 2. La aplicación de back-end tiene como objetivo servir de puente entre la aplicación front-end y los diversos servidores que proveen las métricas de los recursos académicos, concentrando de esta forma la mayor carga de trabajo en sí misma. La aplicación de front-end realiza peticiones a la de back-end mediante una API REST, recopilando los resultados, y es la encargada de generar las visualizaciones en un portal web.

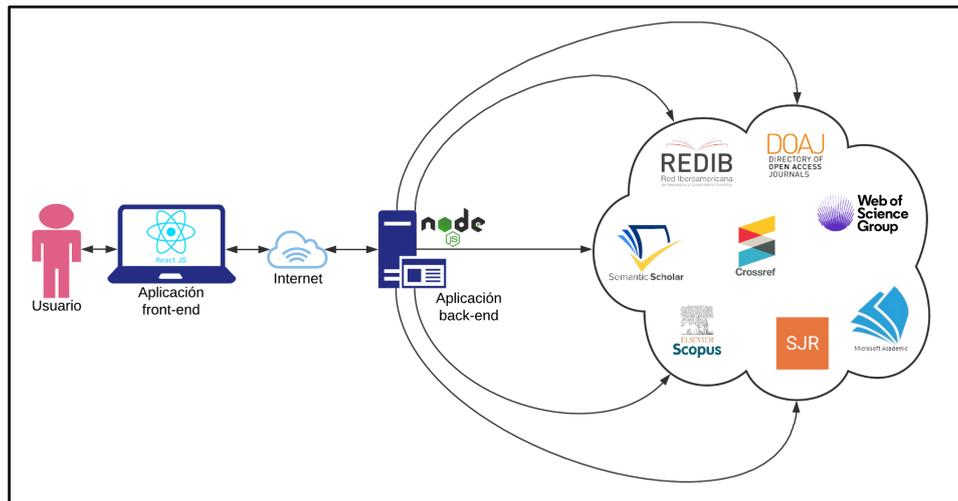


FIGURA 2. Esquema de comunicación entre las aplicaciones

## Descripción de la aplicación back-end

Esta aplicación fue desarrollada con NodeJS y utilizando el framework `express`<sup>2</sup> para definir los endpoints de la API. También incorpora el módulo `node-fetch` para ser capaz de realizar peticiones por HTTP a las diferentes bases de datos y procesar las respuestas. Consta de dos funcionalidades principales:

- Almacenamiento de constantes de acceso, requeridas para acceder a las APIs de Scopus y Microsoft Academic.
- Exposición de una API que brinde la funcionalidad necesaria para que la aplicación de front-end pueda delegar las solicitudes de datos a las diferentes bases de datos.

---

<sup>2</sup> <https://expressjs.com/>

MÉTODO HTTP	ENDPOINT	PARÁMETROS	DESCRIPCIÓN
POST	/api/getContentsDefault	- req: objeto con requerimientos. Se extrae de forma general el campo "url" del body del mismo para indicar a qué API se desea hacer una petición de contenido	Retorna una representación en JSON de la información que responde el sitio web indicado por el campo "req.body.url"
POST	/api/getContentsMicrosoft		Realiza lo mismo que /api/getContentsDefault, agregando la clave de suscripción de Microsoft que está almacenada en esta aplicación
POST	/api/getContentsScopus		Realiza lo mismo que /api/getContentsDefault, agregando la clave de suscripción de Scopus que está almacenada en esta aplicación
POST	/api/getContentsHtml		Retorna un objeto cuyo campo "html" es la representación en formato string del HTML del sitio web "req.body.url". Utilizado para tareas de <i>web scraping</i>

TABLA 2. Endpoints de la API

## Descripción de la aplicación front-end

La aplicación de front-end fue realizada con ReactJS, con el objetivo de elaborar una aplicación web en la que se visualicen las métricas recopiladas. Esta es la encargada de interpretar el pedido de un usuario y comunicarse con la aplicación de back-end para obtener datos de métricas relevantes.

ReactJS es una librería de JavaScript, que permite desarrollar interfaces de usuario interactivas simples. El modelo en el cual se presentan los datos en aplicaciones que utilizan ReactJS está basado en componentes, los cuales encapsulan su propio estado y forma de renderizarse, además de funciones

propias que pudieran tener cada uno<sup>3</sup>. En una página desarrollada con ReactJS, todo lo que se visualiza son componentes, y la filosofía principal es la de no recargar toda la página cuando sólo una porción de la información en ella cambia, sino sólo las partes competentes. Para ello, ReactJS detecta cambios en los estados de sus componentes de manera automática y eficiente, lo que desata re-renderizaciones sólo de aquellos en los cuales el estado es afectado.

Por ser una librería de JavaScript, la parte lógica y funcional de la aplicación puede ser programada con JavaScript puro o incorporar cualquier otra librería o framework que se desee. Para el caso de esta aplicación, en la parte funcional el único framework utilizado es la Web API Fetch<sup>4</sup>, similar al uso de XMLHttpRequest.

## HERA en funcionamiento

### Búsqueda de *journal*

Al introducir un ISSN en la barra de búsqueda, se realiza la búsqueda de un *journal* o publicación periódica. Se obtiene primero la información sobre la editorial y enlace al recurso, así como la vista del resumen de métricas recopiladas de los diferentes sitios, visualizado en la FIGURA 3. Deslizándose hacia abajo se visualiza la vista resumida de las métricas del *journal* que fueron recopiladas en las bases de datos. Contamos con un botón para expandir la información y así visualizar la totalidad de su contenido, como mostramos en las FIGURAS 4, 5 y 6.

---

<sup>3</sup> ReactJS: An Open Source JavaScript Library for Front-end Development - <https://www.theseus.fi/handle/10024/130495>

<sup>4</sup> Fetch API - Web APIs | MDN - [https://developer.mozilla.org/en-US/docs/Web/API/Fetch\\_API](https://developer.mozilla.org/en-US/docs/Web/API/Fetch_API)



FIGURA 3. Resultados generales de la búsqueda de un journal

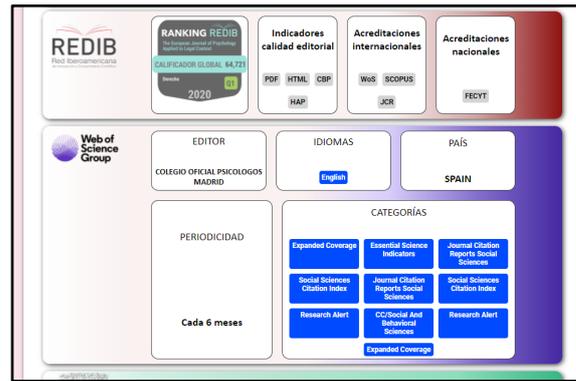


FIGURA 4. Vista expandida de métricas de un journal (1)

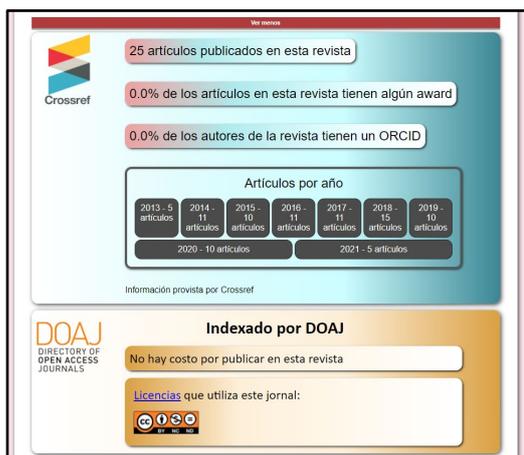


FIGURA 5. Vista expandida de métricas de un journal (2)

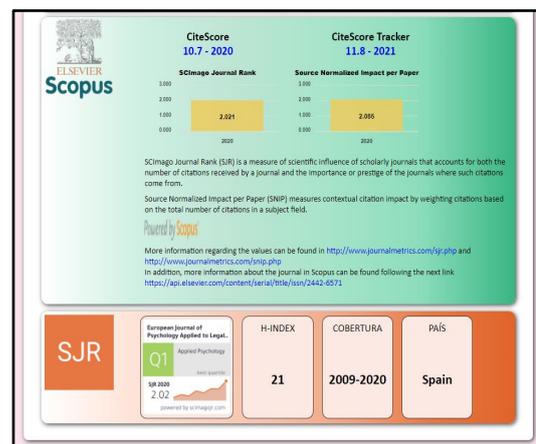


FIGURA 6. Vista expandida de métricas de un journal (3)

## Búsqueda de paper

Al introducir un DOI en la barra de búsqueda y presionar el botón buscar, las métricas son recopiladas de las diferentes bases de datos y se renderizan en pantalla, como se muestra en la FIGURA 7. Entre los resultados se pueden apreciar algunos metadatos de los recursos, como título, tipo de recurso, resumen, entre otros, así como el nombre de la publicación donde se publicó el *paper* y un enlace para dirigirse a su página original. También, de ser posible, la aplicación intentará recolectar métricas del journal donde se encuentra el paper para mostrarlas en conjunto, como se ve en la FIGURA 8. Para ello, se extrae de

los metadatos recuperados el ISSN del journal y se ejecuta de manera automática y en background una búsqueda por ISSN. De esta forma se amplía el contexto de evaluación de un trabajo, al poder visualizar sus métricas de forma aislada y también tener acceso a las del journal y así entender qué influencia podría tener en las del primero. En general, la calidad de un paper está asociada a la del journal en el que está publicado. Al presionar el botón “Ver más” que se ve en esta última imagen, se genera una vista expandida de las métricas en la cual vemos información en forma más detallada (en aquellos casos donde las bases de datos nos provean algún dato adicional), ilustrado en la FIGURA 9.

The screenshot shows the HERA website interface. At the top, there are navigation links for 'HERA', 'Buscar', and 'Preguntas frecuentes'. The main heading is 'Herramienta para Enriquecimiento de Recursos Académicos'. Below this is a search input field with the text 'Ingrese DOI del contenido a analizar' and the value '10.3389/fenvs.2020.561591'. A 'Buscar' button is to the right. The search results display the following information:

- Title:** Analysis of Water Pollution Using Different Physicochemical Parameters: A Study of Yamuna River
- Tipo de recurso:** Journal article
- Autores:** Sharma Rohit, Kumar Raghendra, Satapathy Suresh Chandra, Al-Ansari Nadhir, Singh Krishna Kant, Mahapatra Rajendra Prasad, Agarwal Anuj Kumar, Le Hai Van, Pham Binh Thai.
- Título de la revista:** Frontiers in Environmental Science
- Editorial:** Frontiers Media S.A.
- Año de publicación:** 2020
- ABSTRACT:** The Yamuna river has become one of the most polluted rivers in India as well as in the world because of the high-density population growth and speedy industrialization. The Yamuna river is severely polluted and needs urgent revival. The Yamuna river in Dehradun is polluted due to exceptional tourist activity, poor sewage facilities, and insufficient wastewater management amenities. The measurement of the quality can be done by water quality assessment. In this study, the water quality index has been calculated for the Yamuna river at Dehradun using monthly measurements of 12 physicochemical parameters. Trend forecasting for river water pollution has been performed using different parameters for the years 2020–2024 at Dehradun. The study shows that the values of four parameters namely, Temperature, Total Coliform, TDS, and hardness are increasing yearly, whereas the values of pH and DO are not rising heavily. The considered physicochemical parameters for the study are TDS, Chlorides, Alkalinity, DO, Temperature, COD, BOD, pH, Magnesium, Hardness, Total Coliform, and Calcium. As per the results and trend analysis, the value of total coliform, temperature, and hardness are rising year by year, which is a matter of concern. The values of the considered physicochemical parameters have been monitored using various monitoring stations installed by the Central Pollution Control Board (CPCB), India.

Below the abstract, there is a row of six metric cards:

- Crossref:** 7 citas
- DOAJ (Directory of Open Access Journals):** 7 citas
- Scopus:** Predicciones de citas
- Scilit:** 7 citas
- Sin información de tendencias:** 15
- Se habla del tema:** 1

FIGURA 7. Resultados generales de la búsqueda de un paper (vista resumida)

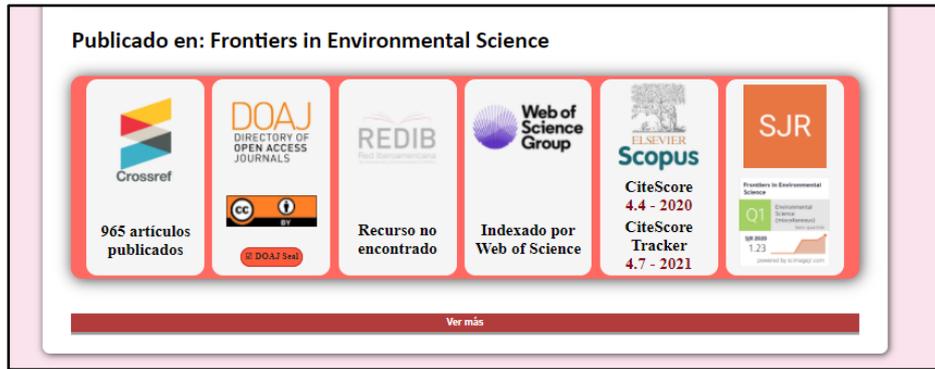


FIGURA 8. Métricas correspondientes al journal donde se encuentra el paper (vista resumida)

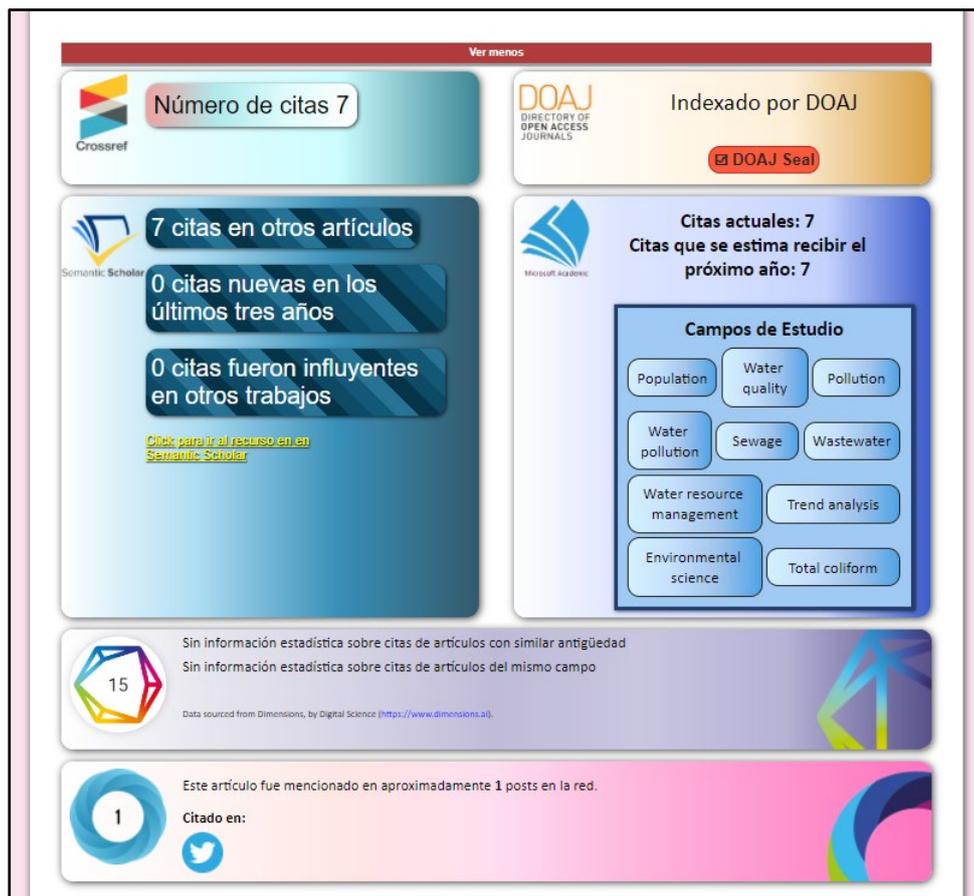


FIGURA 9. Vista expandida de métricas de un paper

## Conclusiones y trabajos futuros

Determinar la calidad y el impacto de una revista o un artículo científico se torna un desafío cada vez más dificultoso debido a dos factores: (1) el crecimiento exponencial en la cantidad de recursos para analizar gracias al desarrollo tecnológico; y (2) la existencia de múltiples indicadores y métricas de diversa índole para cada clase de recurso. En este contexto, HERA representa una herramienta que de manera ágil permite integrar y visualizar métricas de bases de datos relevantes y reconocidas en un único sitio. Esto es de vital importancia a la hora de evaluar contenido, no sólo por eximir a los investigadores de la tediosa tarea de buscar métricas en muchas páginas web, sino también por facilitar la visualización de métricas para su análisis integral.

Considerando las características de HERA, esperamos que los miembros de la comunidad académico-científica la encuentren útil para evaluar la calidad y el impacto de los recursos académicos y que contribuya a facilitar y acelerar dicha tarea.

Por otro lado, el diseño modular de esta aplicación permite y alienta su crecimiento hacia nuevas y mejores funciones para los usuarios. Entre las nuevas funciones que están siendo consideradas, se destacan:

**Expandir el banco de bases de datos académicas:** se buscará incorporar nuevos índices y fuentes de datos en el futuro para ser añadidos en la herramienta. Opciones atractivas incluyen la incorporación de índices regionales, como por ejemplo Redalyc, y también bases reconocidas pero de disciplinas más específicas como el caso de PubMed. Dada esta posibilidad de contar con numerosas bases, sería conveniente en ese caso incorporar la funcionalidad de poder seleccionar de qué bases de datos interesa obtener información.

**Implementación de una API REST:** la herramienta podría exponer parte de su funcionalidad mediante el desarrollo de una API REST para así desarrollar nuevas herramientas que se nutran de los datos brindados por HERA. Por ejemplo, de tener una API se podrían implementar soluciones como:

- Extensiones web para navegadores que muestren métricas de un recurso directamente desde la página de dicho recurso.
- Generación de archivos de datos que recopilen métricas de varios recursos académicos a la vez. Esta funcionalidad puede resultar ideal para evaluar recursos dentro de un contexto específico, como un proyecto particular, un área temática o una institución.
- Integración con OJS para visualizar métricas desde allí y obtener datos a través de terceros de forma más transparente.

**Optimización de tiempo de búsquedas de métricas:** sería de utilidad poseer mecanismos de caché para optimizar los tiempos en caso de búsquedas recurrentes. Además, algunas de las API de sitios externos utilizadas tienen cuotas máximas específicas de uso, que se podría evitar alcanzar al tener la información ya almacenada.

## Bibliografía

- GONZALEZ-PARDO, R.; REPISO, R.; ARROYAVE-CABRERA, J. (2020). Revistas iberoamericanas de comunicación a través de las bases de datos Latindex, Dialnet, DOAJ, Scopus, AHCI, SSCI, REDIB, MIAR, ESCI y Google Scholar Metrics. *Revista Española de Documentación Científica*, 43 (4): e276. <https://doi.org/10.3989/redc.2020.4.1732>
- LINDSEY, D. (1989). Using citation counts as a measure of quality in science measuring what's measurable rather than what's valid. *Scientometrics*, 15 (3-4): 189-203.
- REPISO, R. (2015). Cómo identificar una revista de calidad. *Cardiocre*, 50 (2): 46-48. <https://www.redalyc.org/articulo.oa?id=277041630002>
- RHAIEM, M. (2017). Measurement and determinants of academic research efficiency: a systematic review of the evidence. *Scientometrics*, 110: 581-615. <https://doi.org/10.1007/s11192-016-2173-1>

Eje temático:

Infraestructura tecnológica

[Ver bloque de ponencias en You Tube](#)

# SUM-OJS: Scripts para actualizaciones y migraciones de OJS

## **GONZALO LUJÁN VILLARREAL**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[gonzalo@prebi.unlp.edu.ar](mailto:gonzalo@prebi.unlp.edu.ar)

## **PABLO GABRIEL TERRONE**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

[pabloterrone@sedici.unlp.edu.ar](mailto:pabloterrone@sedici.unlp.edu.ar)

## **MARISA RAQUEL DE GIUSTI**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[marisa.degiusti@sedici.unlp.edu.ar](mailto:marisa.degiusti@sedici.unlp.edu.ar)

## **RESUMEN**

En este trabajo presentamos SUM-OJS, una herramienta desarrollada con el objetivo de acelerar y simplificar las tareas vinculadas a actualizaciones del software Open Journals System (OJS) hacia nuevas versiones, así como también a migraciones hacia nuevos entornos de ejecución. Esta herramienta utiliza una estructura de proyectos basada en Docker y Docker-compose, que permite generar el entorno de ejecución, actualización o migración requerido por cada versión particular de OJS, y una serie de comandos que automatizan muchas de las tareas recurrentes que se deben ejecutar durante las migraciones. Esta ponencia explica la arquitectura interna y forma de uso de SUM-OJS, e incluye un listado de actualizaciones de OJS

realizadas con esta herramienta, junto a los problemas identificados y soluciones aplicadas en cada caso.

**PALABRAS CLAVE**

Open Journals System, actualización, migración, virtualización.

## **SUM-OJS: Scripts for updates and migrations of OJS**

**ABSTRACT**

In this paper we propose SUM-OJS, a software tool developed to accelerate and simplify common tasks while running upgrades of Open Journals System (OJS), as well as testing migrations towards new environments. This software has been built upon a Docker and Docker-compose project structure, which enables a fast and custom setup of different environments for running, updating or migrating required for each particular OJS version. We have also included a command line tool to run common tasks during migrations. The paper introduces the inner architecture of SUM-OJS and a quick hands-on tutorial, and it also includes a list of upgrades already performed with this tool, with many issues identified and solutions applied in each case.

**KEYWORDS**

Open Journals System, upgrade, migration, virtualization.

## **Introducción**

Open Journals System (OJS) es un sistema de gestión de publicaciones periódicas de código abierto. Su desarrollo es liderado por el Public Knowledge Project (PKP), y la primera versión fue publicada en el año 2005. Desde entonces, el trabajo mancomunado entre PKP y una creciente comunidad de

desarrolladores en todo el mundo ha permitido incorporar cada vez más funciones, optimizar las ya existentes y modernizar las librerías y herramientas sobre las que funciona OJS. Las actualizaciones de OJS se publican de manera periódica, y con cada nueva actualización se incorporan nuevas funcionalidades y se resuelven problemas reportados en versiones anteriores, entre los que se incluyen cuestiones vinculadas a la seguridad, a la performance y uso de recursos, y a fallas en el funcionamiento de determinadas áreas del sistema. En muchas ocasiones, los cambios de versión también requieren realizar cambios en los servicios que utiliza OJS: versión de PHP (por ejemplo de PHP 5.6 a 7.0 y luego a 7.3), motor de almacenamiento en MySQL (por ej. de MyISAM a InnoDB), instalación de librerías adicionales en alguno de los servicios, asignación de memoria, entre otros.

Las actualizaciones de OJS pueden clasificarse en dos categorías: actualizaciones menores, como la actualización desde OJS 3.2.1 hacia OJS 3.2.1-1 y luego hacia 3.2.1-2, que corrigen errores o problemas puntuales, y actualizaciones mayores, por ejemplo desde OJS 3.0 a OJS 3.1 y luego a OJS 3.2, que incorporan un gran número de mejoras y en muchos casos realizan cambios a nivel de arquitectura: cambio de framework MVC, nuevo sistema de gestión de traducciones, cambios en la jerarquía de clases de plugins y/o temas, etc.

Entre las distintas actualizaciones de OJS, puede establecerse un punto de inflexión en el año 2015, con el lanzamiento de OJS 3.0. Esta versión propuso cambios globales en el sistema, que si bien mantenía su esencia en cuanto a flujos de trabajo, roles y servicios esenciales, introducía una interfaz de usuario completamente rediseñada, junto a un nuevo *framework* desarrollado por PKP para todos sus sistemas, y una gran cantidad de opciones de configuración y personalización que otorgaban a OJS mayor flexibilidad para adecuarse a los requerimientos de las distintas instituciones que lo utilizan. Desde el lanzamiento de OJS 3.0, se han generado más de treinta actualizaciones, cuatro de ellas mayores (3.0, 3.1, 3.2 y 3.3); si bien el lanzamiento de cada nueva versión es coordinado por el propio desarrollador PKP, es importante

destacar el rol de un gran número de instituciones académicas, empresas editoriales y usuarios independientes de todo el mundo, quienes realizan sus aportes con código fuente, reportes de errores y traducciones.

Con cada nueva versión, el equipo de PKP empaqueta junto al código fuente una herramienta (*script*) que ejecuta las tareas requeridas para realizar la actualización de la base de datos. Las tareas que realiza este script se encuentran en la incorporación de nuevas opciones de configuración, la creación o modificación de tablas, la modificación de registros tanto por correcciones de errores como por modificaciones en el modelo de datos, el cambio de nombre de archivos y/o directorios, entre otras. Asimismo, antes de ejecutar las tareas de actualización, el script realiza algunos controles vinculados al sistema subyacente, como por ejemplo que la versión de PHP sea adecuada, o que se cuente con los permisos necesarios sobre determinados directorios y archivos.

Una característica interesante de esta herramienta para la actualización es su metodología incremental, que ejecuta los cambios de manera iterativa desde la versión instalada hacia la versión destino. Esto significa que, si una organización posee la versión 3.0 de OJS, y desea pasar a la versión 3.2, la herramienta se encargará de llevar primero desde la versión 3.0 a la versión 3.1, luego de la versión 3.1 a la versión 3.2. En condiciones normales, esta característica resulta muy conveniente, ya que ahorra mucho trabajo, dado que se evita realizar actualizaciones manuales entre las versiones intermedias.

Al momento de realizar una actualización, el proceso puede demorar desde pocos minutos hasta algunas horas. El tamaño de la base de datos, relacionado con la cantidad de revistas, de envíos y de usuarios, así como la capacidad de procesamiento (CPU, memoria, discos), son los factores que más influyen en el tiempo total del proceso de actualización. Una vez finalizado dicho proceso, y en caso de que no se hayan observado errores, deberán revisarse las nuevas características y funcionalidades de la versión de OJS actualizada, habilitar o re-configurar plugins, y realizar otras tareas de mantenimiento y verificación antes de habilitar el uso o el acceso a los usuarios del sistema. Sin embargo, en

caso de detectarse un problema durante la actualización, el script informará al usuario sobre las incidencias encontradas, y queda a criterio de cada usuario decidir la mejor estrategia para corregir el error y volver a intentar la actualización. Este proceso puede requerir numerosas iteraciones, ya que la corrección de un problema puede exponer un segundo problema que antes no era visible. Es difícil determinar el porcentaje de actualizaciones que fallan: en los foros de PKP<sup>1</sup> se observan una cantidad importante de hilos donde se reportan errores de todo tipo, gracias a la participación de la comunidad de usuarios y desarrolladores en la búsqueda de soluciones, que podrían incluso generar cambios en futuras versiones de OJS para corregir los errores detectados.

Para evitar problemas vinculados a la pérdida de información, a la generación de datos inconsistentes o a la imposibilidad de acceder a archivos alojados en el servidor, es conveniente realizar las actualizaciones en un entorno de desarrollo o pruebas, aislado e independiente del servidor en producción. Asimismo, como se mencionó previamente, al identificar un problema en la actualización se deberán realizar las acciones correctivas para resolver dicho problema, para luego volver a intentar la actualización. Sumado a esto, no siempre es posible actualizar directamente hacia la última versión, ya que los errores pueden surgir en alguna etapa intermedia del proceso de actualización. En estos casos, se deben realizar actualizaciones parciales, y realizar correcciones intermedias antes de continuar la actualización. El Gráfico 1 ilustra de manera simplificada este proceso para actualizar desde la versión X a la Z, pasando por la versión intermedia Y.

Luego de realizar muchas actualizaciones entre versiones de OJS, los autores de este trabajo hemos buscado formas de automatizar algunas de las tareas recurrentes, minimizar la tasa de errores y organizar todas las tecnologías (en sus diferentes versiones) requeridas. A partir de aquí surge

---

<sup>1</sup> Hilos en los foros de PKP vinculados problemas de actualización de OJS: <https://forum.pkp.sfu.ca/tags/c/questions/5/upgrade>

SUM OJS, una herramienta que facilita la repetición de este proceso de actualización de OJS a partir de diferentes estrategias, como la definición de contextos de ejecución particulares, replicación de entornos para pruebas y automatización de tareas de corrección de errores. Como se indica en el apartado “Casos de uso en contextos reales”, esta herramienta ha sido puesta a prueba en numerosas instalaciones de OJS, que fueron actualizadas a las diferentes versiones disponibles al momento de realizar la actualización, con resultados satisfactorios.

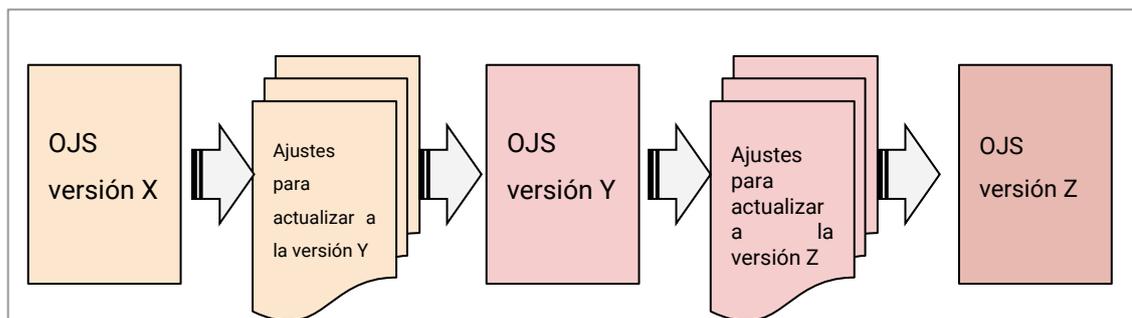


GRÁFICO 1. Etapas en la actualización de OJS desde la versión X a la versión Z ( $X < Y < Z$ ) (autoría propia)

## Desarrollo

Como se mencionó previamente, el proceso de actualización puede demorar mucho tiempo si se trata de grandes instalaciones de OJS. Este proceso se realiza de manera iterativa entre versión y versión, y cada intento de actualización requerirá, como mínimo:

1. Preparar el servidor para ejecutar la actualización (servidor web, versión de PHP, servidor de base de datos).
2. Cargar una base de datos con la versión de OJS en curso.
3. Copiar los archivos subidos por los usuarios del OJS en curso.
4. Realizar los ajustes necesarios para poder pasar a la siguiente versión.
5. Ejecutar la actualización hacia la siguiente versión.

6. **a.** Si la actualización del paso previo falló, revisar la causa, preparar las correcciones para realizar un nuevo intento y volver al paso 1. **b.** Si la actualización del paso previo funcionó, continuar con la actualización hacia la siguiente versión.

Si bien este proceso describe a muy alto nivel las distintas tareas involucradas en cada etapa, cabe destacar que algunas tareas pueden demorar mucho tiempo por su complejidad o volumen de información, y pueden introducir nuevos errores, en especial en los casos que requieran configuraciones específicas en alguno de los servicios del sistema (PHP, servidor web o servidor de base de datos).

Para agilizar este proceso iterativo, hemos implementado un *stack* de ejecución basado en *containers dockers*, junto a unos *scripts bash* y *Makefile*. El uso de los contenedores permite automatizar o agilizar algunas de las tareas descritas previamente. A continuación, se brinda una breve introducción a estas tecnologías, y en la siguiente sección se describen los componentes de este *stack*, indicando el rol de cada uno de ellos en proceso de ejecución de actualizaciones y migraciones.

## Virtualización basada en contenedores Docker

Docker es una herramienta cada vez más utilizada por la comunidad científica, gracias a su capacidad de ejecutar aplicaciones auto-empaquetadas en entornos aislados<sup>2</sup>. Una de las ventajas de Docker radica en su portabilidad, que le permite distribuir una misma aplicación en un amplio rango de plataformas (DI TOMMASO, 2015).

---

<sup>2</sup> Docker: <https://www.docker.com/>

A diferencia de la forma de virtualización completa, en la cual se virtualiza el hardware y se levanta un sistema operativo con todas sus funcionalidades para ejecutar el servicio particular, Docker realiza una virtualización mucho más liviana ya que no virtualiza el hardware y utiliza el núcleo del sistema operativo del host dentro de los contenedores.

Una de las ventajas más evidentes a la hora de utilizar Docker es la habilidad de descargar imágenes pre-construidas que contienen paquetes de software y sus dependencias listos para usarse, evitando así su instalación y configuración manual por parte de los usuarios. Otra ventaja no menos importante es que se ejecuta cada proceso en un contenedor aislado, lo cual evita problemas con otros servicios y asegura que se ejecutará en una configuración predecible, que no cambiará a lo largo del tiempo debido a actualizaciones ni errores de configuración del software.

## Organización de servicios con Docker Compose

Cuando se desea ejecutar muchos servicios para crear un entorno de desarrollo, resulta fundamental poder organizarlos. Docker Compose (IBRAHIM, 2021) permite, en un solo archivo, definir un conjunto de servicios que trabajan de forma colaborativa en un entorno aislado. Para esto permite definir aspectos como:

- Mapeo de puertos: cada servicio puede mapear los puertos virtuales que utilizará con los puertos físicos.
- Establecer dependencias: permite indicar el orden de ejecución en los servicios, lo que es especialmente útil cuando un servicio depende de otro para poder funcionar.
- Mapeo de volúmenes: permite definir volúmenes para montar dentro de containers, en el contexto de este trabajo se aprovecha la posibilidad de poder mapear volúmenes con directorios del sistema operativo host.

- Utilización de variables de entorno: muchos servicios necesitan ciertos datos para poder ejecutarse (host, nombre de usuario, contraseña, etc), Docker Compose permite que estos datos se recuperen desde un archivo externo, por lo que resulta fácil cambiar alguno de los datos.

## Bash y Makefile

Cuando se trabaja con procesos iterativos, que poseen muchas instrucciones que pueden repetirse más de una vez, es conveniente diseñar mecanismos para automatizar estos procesos. Hacer estas repeticiones de forma manual (copiando y pegando las instrucciones) puede ser problemático ya que algunas veces pueden cometerse errores sintácticos u omitirse alguna instrucción (BAKER, 2020). Sumado a esto, a medida que aumenta la cantidad de instrucciones a ejecutarse, su ejecución manual requerirá cada vez más tiempo. Para evitar estos inconvenientes en este proyecto utilizamos las tecnologías Bash y GNU Make.

Bash, un lenguaje de *scripting*, muy utilizado para la automatización de tareas en entornos GNU/Linux, posee una amplia variedad de instrucciones altamente parametrizables que pueden agruparse en un único archivo (BASH, 2022), lo cual resulta muy útil lograr procesos de migración mucho más rápido y sencillo.

GNU Make utiliza archivos Makefile, como punto de entrada a tareas complejas, agrupando conjuntos de instrucciones o *scripts* en una sola instrucción *make* (MAKE, 2022). En nuestro caso, lo utilizamos para simplificar la invocación a instrucciones de Docker y a los scripts de Bash, generando así código más fácil de escribir, mantener, resultando así en programas más fáciles de utilizar.

A modo de ejemplo, tomamos uno de los comandos incluidos en el Makefile, correspondiente con *make regenerate*. Dicho comando invoca al *script bash upgrader*, al cual le envía los parámetros correspondientes que incluyen la operación a realizar (parámetro -o), la versión desde la que se parte (parámetro

-f) y la versión a la que se desea actualizar (parámetro -t). Todo esto se resume en la siguiente línea:

```
./upgrader -o regenerate -f $(from) -t $(to)
```

Internamente, el *upgrader* ejecuta diversas instrucciones para hacer lo siguiente:

1. Elimina la base de datos actual.
2. Crea una nueva base de datos vacía.
3. Carga el backup de la base de datos.
4. Prepara la base de datos para la migración, ejecutando las consultas del `post_dump_file`.

De esta manera el proceso de migración es mucho más sencillo, se resume en levantar la base de datos (*Make up*), regenerarla (*Make regenerate*) y ejecutar la actualización (*Make upgrade*).

## Descripción del *stack*

Como se describió previamente, basados en la experiencia personal, la actualización entre versiones de OJS es un proceso iterativo, que requiere realizar pruebas y ajustes entre las pruebas a fin de alcanzar una actualización *limpia*: sin errores ni durante el proceso de actualización, ni durante las pruebas posteriores a la actualización. Estos ajustes y correcciones pueden clasificarse en dos grandes categorías:

- **Ajustes a nivel de servidor:** aquí se agrupan los cambios entre versiones de PHP, en la asignación de límites de memoria, configuración de encabezados HTTP en el servidor web, cambios en el tamaño de los buffers del servidor de bases de datos, entre otros.

- **Ajustes a nivel de base de datos:** involucra todas las correcciones que se hacen sobre la base de datos de OJS, lo que abarca la eliminación de valores nulos, la corrección de referencias incorrectas, la corrección de problemas de codificación de caracteres, entre otros.

Para agilizar la aplicación de ajustes a nivel del servidor, se decidió por una solución basada en contenedores Docker, organizados en una serie de pilas de ejecución implementadas sobre Docker Compose<sup>3</sup> que incluyen al servidor de bases de datos, al servidor web con PHP, y al software phpMyAdmin para realizar tareas de verificación de datos, búsqueda de errores, etc. La elección de la tecnología de contenerización basada en Docker asegura un modelo de ejecución de *containers* flexible y altamente configurable, que permite emular entornos de desarrollo y ejecución específicos en poco tiempo. Además, se trata de herramientas muy utilizadas, con una gran comunidad de usuarios y una amplia gama de servicios disponibles para los desarrolladores: repositorios de imágenes, documentación, servicios de monitoreo, etc.

A continuación, se muestra un extracto de uno de los archivos docker-compose, en el que se muestran los principales elementos de la configuración:

```
version: '3.7'

services:

  mysql:

    image: mysql:5

    container_name: ${PROJECT_NAME}_mysql

  volumes:

    - ./sql:/var/backups

  environment:
```

---

<sup>3</sup> Compose specifications, Docker documentation:  
<https://docs.docker.com/compose/compose-file/>

```
MYSQL_DATABASE: ${PROJECT_NAME}

phpmyadmin:
  depends_on:
    - mysql
  image: phpmyadmin/phpmyadmin
  container_name: ${PROJECT_NAME}_phpmyadmin
  environment:
  ports:
    - '8080:80'

web:
  #for OJS 3.2
  image: webdevops/php-apache-dev:7.3
  container_name: ${PROJECT_NAME}_web
  volumes:
    - "./public_ojs:/app"
    - "../private_ojs:/var/ojs-data/"
  ports:
    - "80:80"
  environment:
    PHP_MEMORY_LIMIT: 512M
    PHP_DISPLAY_ERRORS: 1

EXTRACTO DE CÓDIGO 1. Configuración de servicios de Docker Compose (autoría propia)
```

La configuración de Docker Compose describe tres servicios: MySQL (motor de base de datos), phpMyAdmin (gestión gráfico de la base de datos) y web (servidor Apache y php). El servicio MySQL está basado en la versión 5, y utiliza un volumen externo (directorio ./sql). El servicio phpMyAdmin se conecta con el servicio MySQL y queda accesible a través del puerto 8080 del host. El servicio web utiliza el servidor web Apache, y el intérprete de PHP versión 7.3; este

servicio también utiliza volúmenes externos, uno donde se encuentra el código fuente de OJS (public\_ojs) y otro donde se encuentran los archivos subidos por los usuarios (private\_ojs). Adicionalmente, el servicio web expone el acceso a OJS a través del puerto 80 del host, e incluye algunas opciones de configuración específicas (512 MB de límite de memoria, y mostrar los errores de php).

## Estructura de directorios

La solución propuesta utiliza una estructura de directorios que contiene por un lado los archivos generales del proyecto y, por otro, subdirectorios con archivos particulares para migrar entre distintas versiones de OJS. A continuación, se incluye un extracto de la estructura de directorios propuesta:

```

— .env
  |— from3.1-to-3.2
  |   |— config.inc.php
  |   |— docker-compose.yml
  |   |— ojs-3.2.1-3.tar.gz
  |   |— public_ojs
  |   |— sql
  |       |— load-data.sql
  |       |— post-load-data.sql
  |— from3.2-to-3.3
  |   |— docker-compose.yml
  |   |— config.inc.php
  |   |— ojs-3.3.0-9.tar.gz
  |   |— public_ojs
  |   |— sql

```

```
|   |— load-data.sql
|   |— post-load-data.sql
|— Makefile
|— private_ojs
|— README.md
|— upgrader
```

LISTADO 1. Estructura de directorios de SUM-OJS (autoría propia)

En esta estructura se observa:

- Un directorio 'private\_ojs', que se utiliza para alojar los archivos subidos por los usuarios de OJS.
- Dos directorios ('from3.1-to-3.2' y 'from3.2-to-3.3') que poseen los archivos y directorios necesarios para realizar migraciones entre las versiones 3.1 a 3.2 y 3.2 a 3.3 respectivamente. Dentro de estos directorios se destacan los siguientes objetos:
  - Un archivo 'docker-compose.yml' con la pila de Docker adecuada para la versión de OJS **hacia** la que se está migrando.
  - Un archivo 'config.inc.php', que posee las configuraciones básicas de OJS para la versión **hacia** la que se está migrando.
  - Un directorio SQL, que posee dos archivos:
    - 'load-data.sql', que posee una copia completa de la base de datos del OJS **desde** el que se está migrando;
    - 'post-load-data.sql', que posee el código SQL para ejecutarse sobre la base de datos previa a migrar.
  - Un directorio 'public\_ojs', donde debería incluirse el sistema OJS en la versión **hacia** la que se está migrando.
- Un archivo ejecutable llamado *upgrader*, que incluye varias rutinas para utilizar los distintos servicios Docker.
- Un archivo *Makefile*, que ofrece una interfaz de línea de comandos de alto nivel para utilizar las rutinas del programa *upgrader*.

- Un archivo '.env' que posee variables de entorno globales (nombre de base de datos, nombres de archivos utilizados en las distintas etapas, usuario y clave de conexión a la base de datos, etc.)

## Modo de uso

Para utilizar este desarrollo, se requiere realizar una configuración inicial por única vez, en la que se preparará todo el entorno para realizar las pruebas de migraciones.

## Aspectos a tener en cuenta durante la configuración inicial global

El directorio 'private\_ojs' debe contener todos los archivos privados de OJS. Se deberán brindar permisos de lectura y escritura de manera recursiva a dicho directorio, para que las migraciones puedan renombrar y mover archivos entre directorios. Además, el archivo '.env' debe contener las variables de entorno con las que se realizará la migración. Estas variables de entorno son:

#Nombre del proyecto, utilizado para crear contenedores docker. Por ejemplo:  
PROJECT\_NAME=ojs\_migrate

#Nombre de la base de datos sobre la que se realizarán las migraciones. Por ejemplo:  
DATABASE\_NAME=ojs\_db

#Contraseña del usuario root, necesaria para regenerar la base de datos cuando se solicita.  
Por ejemplo: ROOT\_PASSWORD=root

#Nombre del archivo SQL que contiene la base de datos a importar cada vez que se genere. Por ejemplo: DUMP\_FILE=load-data.sql

#Archivo SQL que contiene otros comandos SQL a ejecutar luego de regenerar la base de datos. Por ejemplo: POST\_DUMP\_FILE=post-load-data-sql

#Directorio dentro de los contenedores donde se alojarán los archivos SQL. Por ejemplo: DUMP\_DIR=/var/backups

EXTRACTO DE CÓDIGO 2. Configuración de variables de entorno de SUM-OJS (autoría propia)

## Aspectos a tener en cuenta para configurar y ejecutar una actualización

Para realizar una actualización desde la versión X a la versión Y de OJS, se debe localizar el directorio 'fromX-to-Y'. Por ejemplo, para realizar la actualización desde la versión 3.1 (X) hacia la versión 3.2 (Y), deberá utilizarse el directorio 'from3.1-to-3.2'. El proyecto posee directorios para realizar actualizaciones entre dos versiones (3.1 a 3.2, y 3.2 a 3.3), pero el esquema propuesto puede extenderse fácilmente para realizar actualizaciones entre cualquier par de versiones, tanto anteriores como posteriores a las ya incluidas.

Antes de realizar la actualización, este directorio debe contener los archivos listos para generar la base de datos y ejecutar la actualización. Para su correcta ejecución, debe:

- Copiar en el directorio from-X-to-Y/sql el archivo con el dump completo de la base de datos del OJS versión X. El nombre del archivo debe ser el especificado previamente en la variable DUMP\_FILE del archivo de configuración global .env.
- En caso de ser necesario, se deberá copiar en el directorio from-X-to-Y/sql el archivo que contenga los comandos SQL a ejecutarse luego de importar la base de datos original, y antes de ejecutar la actualización hacia la versión Y. El nombre de este archivo debe ser el especificado previamente en la variable POST\_DUMP\_FILE.

- Asegurarse de que el *stack* definido en el archivo “from-X-to-Y/docker-compose.yml” cumple con los requerimientos para ejecutar OJS versión Y. En particular, debe observarse que la versión de PHP sea adecuada en la línea “image: webdevops/php-apache-dev:PHP\_VERSION”. Por ejemplo, para ejecutar una actualización hacia OJS 3.2, puede utilizarse PHP 7.3, especificada de la siguiente forma: image: webdevops/php-apache-dev:7.3.
- Descomprimir en el directorio ‘from-X-to-Y/public\_ojs’ una copia de OJS versión Y.
- Asignar permisos de lectura y escritura sobre el directorio de cache de OJS: `chmod a+rw from-X-to-Y/public_ojs/cache -R`.
- Generar un archivo de configuración para OJS versión Y (puede realizarse a partir del archivo TEMPLATE incluido en OJS). Este archivo debe llamarse ‘config.inc.php’, debe estar alojado dentro del directorio ‘public\_ojs’, y deberá incluir la configuración necesaria para que OJS versión Y se ejecute desde el contenedor Docker. A continuación, se incluyen algunas de las opciones de configuración más relevantes que deberán tenerse en cuenta:

```
base_url = "http://localhost"
```

You might also need to set override settings, such as:

```
base_url[SOME_JOURNAL] = http://localhost/SOME_JOURNAL
```

```
[database]
```

```
driver = mysqli
```

```
host = mysql
```

```
username = root
```

```
password = <--- same as password set in .env
```

```
name = <--- same as database set in .env
```

```
[files]
```

```
files_dir = /var/ojs-data/uploads
```

```
public_files_dir = public
```

EXTRACTO DE CÓDIGO 3. Configuración de la instalación de OJS (autoría propia)

Una vez realizado esto, se cuenta con un entorno de ejecución de OJS versión Y, y con una copia de la base de datos de OJS versión X. Ahora es hora de utilizar los scripts de automatización de tareas para regenerar la base de datos en versión X, así como también para ejecutar las migraciones.

El formato de los comandos es el siguiente:

**make** *operación* **from=X to=Y**.

Las operaciones disponibles son:

- **up**: inicia los contenedores docker para migrar OJS desde la versión X hacia la versión Y.
- **down**: detiene los contenedores docker levantados previamente con up.
- **regenerate**: regenera la base de datos de la versión X, y ejecuta los scripts SQL previos a la actualización hacia la versión Y.
- **upgrade**: ejecuta los scripts de actualización hacia la versión Y.

Por ejemplo, para realizar la migración desde OJS 3.1 a OJS 3.2, los comandos a ejecutar serán:

```
make up from=3.1 to=3.2
```

```
make regenerate from=3.1 to=3.2
```

```
make upgrade from=3.1 to=3.2
```

```
make down from=3.1 to=3.2
```

En caso de que la migración falle, será necesario volver a generar la base de datos (utilizando el comando `make regenerate from=X to=Y`). Esto eliminará la base de datos actual y la volverá a cargar a partir del backup. Una vez que se restauró, se deberán realizar las correcciones que provocaron el fallo en el intento anterior. Si estas correcciones requieren modificar la base de datos (por ejemplo: eliminar nulls, revisar codificación, etc.), estos comandos deberían incorporarse al archivo `POST_DUMP_FILE` a fin de automatizar su ejecución en futuras pruebas. Por el otro lado, si las correcciones implican cambios a nivel de la infraestructura (versión de PHP, límites de memoria, configuraciones de MySQL), entonces los cambios deberán aplicarse sobre el archivo `'docker-compose'`.

Luego, realizadas estas correcciones, se puede volver a intentar la actualización ejecutando la misma secuencia de comandos. Una vez que la actualización ha sido exitosa, ya se contará con OJS ver. 3.2 actualizado, y se podrá actualizar hacia la siguiente versión. Sin embargo, en este punto sería deseable realizar una copia de seguridad de la base de datos, ya que, en caso de que la actualización hacia la versión 3.3 falle, podrá volver a intentarse desde la versión 3.2, en vez de volver a intentarlo desde la versión 3.1.

Para realizar una copia de seguridad, puede utilizarse el servicio phpMyAdmin (incluido en `'docker-compose'`) a través de un navegador web en la dirección (<http://localhost:8080>), o también puede realizarse desde la línea de comandos, lo que puede resultar más eficiente en muchos casos. Para ello, se ha incluido un comando que realiza estas copias de seguridad y las almacena en el directorio `DUMP_DIR`, junto a otros comandos que permiten iniciar líneas de comandos en los contenedores MySQL y WEB, a fin de ejecutar operaciones rápidas desde la línea de comandos.

- **make backup**: realiza un backup en el directorio `DUMP_DIR`.
- **make bash-mysql**: inicia una terminal dentro del contenedor `mysql`.
- **make bash-php**: inicia una terminal dentro del contenedor `web`.

## Casos de uso en contextos reales

Para verificar el funcionamiento de esta herramienta se realizaron migraciones de distintas instalaciones de OJS. La siguiente tabla resume en orden cronológico (desde mediados de 2019 a principios de 2022) los trabajos de actualización realizados:

REVISTA	DESCRIPCIÓN DE LA INSTALACIÓN	VERSIÓN ORIGEN	VERSIÓN DESTINO
<i>Ameghiniana</i> <a href="https://www.ameghiniana.org.ar/">https://www.ameghiniana.org.ar/</a>	Instalación de OJS con una única revista. Idioma inglés	2.4.8-3	3.2.1-1
<i>PE-APA</i> <a href="https://www.peapaleontologica.org.ar/">https://www.peapaleontologica.org.ar/</a>	Instalación de OJS con una única revista. Idiomas inglés y español	2.4.8-3	3.2.1-1
<i>Revista de la Asociación Geológica Argentina</i> <a href="https://revista.geologica.org.ar/">https://revista.geologica.org.ar</a>	Instalación de OJS con una única revista. Idioma español	2.3.6-0	3.2.1-1
Revistas de la Sociedad Argentina para el Estudio de los Mamíferos <a href="https://ojs.sarem.org.ar/">https://ojs.sarem.org.ar/</a>	Instalación de OJS con dos revistas. Idiomas inglés y español	2.4.8-1	3.3.0-6
<i>Revista de Geología Aplicada a la Ingeniería y al Ambiente</i> <a href="https://www.editoresasagai.org.ar/ojs/">https://www.editoresasagai.org.ar/ojs/</a>	Instalación de OJS con una única revista. Idioma español	2.4.8.1	3.3.0-7
<i>PE-APA</i> (segunda actualización) <a href="https://www.peapaleontologica.org.ar/">https://www.peapaleontologica.org.ar/</a>	Instalación de OJS con una única revista. Idiomas inglés y español	3.2.1-1	3.3.0-8
Portal de Revistas de la UNLP <a href="https://revistas.unlp.edu.ar/index">https://revistas.unlp.edu.ar/index</a>	Instalación de OJS con múltiples revistas. Múltiples idiomas	3.1.2-1	3.3.0-9

TABLA 1. Actualizaciones a instalaciones de OJS utilizando SUM-OJS (autoría propia)

En líneas generales, el proceso de migración se realizó de la siguiente manera:

- Para instalaciones previas a OJS 2.4, se realizó una actualización hacia OJS 2.4.8-4. Para ello, se generó un *stack* con PHP 5.6 y MySQL 5.5. Cabe destacar que se utilizó la librería PDO para conectar al servicio MySQL desde PHP.
- Para instalaciones 2.4.x, se realizó primero una actualización hacia OJS 3.0. Para ello, se generó un *stack* con PHP 7.0 y MySQL 5.5. Aquí también se utilizó PDO.
- Para instalaciones 3.0.x, se realizó la actualización hacia 3.1. Para ello se generó un *stack* con PHP 7.2 y MySQL versión 8 (o MariaDB versión 10 o superior). A partir de aquí se comienza a utilizar el conector `mysqli`.
- Para instalaciones 3.1, se realizaron actualizaciones a OJS 3.2. Para ello se generó un *stack* con PHP 7.3 ó 7.4 (indistintamente), con MySQL 8/MariaDB 10 y `mysqli`.
- Para instalaciones 3.2, se realizaron actualizaciones hacia OJS 3.3. Para ello se generó un *stack* con PHP 7.4, con MySQL 8/MariaDB 10 y `mysqli`.

Como puede observarse, las actualizaciones entre distintas versiones requieren configuraciones y combinaciones de servicios particulares, lo cual fue posible de realizar de manera muy ágil gracias a esta propuesta basada en Docker y Docker-compose. Asimismo, además de los distintos *stacks* generados para cada versión de OJS, es importante destacar que cada actualización requirió sus propias tareas para la corrección de errores, limpieza de datos, ajustes de configuraciones, etc. A lo largo de las distintas actualizaciones, las necesidades particulares de cada caso requirieron ajustes en la herramienta aquí presentada, en particular orientadas a automatizar tareas recurrentes (como por ejemplo regeneración de la base de datos, acceso a la línea de comandos del servicio web o creación de backups).

Si bien existen muchos motivos por los cuales una actualización de OJS puede fallar, luego de aplicar el proceso de actualización en distintos contextos se identificaron algunos factores que determinan la cantidad de “trabajo extra” (ajustes, limpieza de datos, entre otros.) que será necesario para obtener una actualización exitosa:

- 1. Volumen de información:** se observó que las instalaciones de OJS con mayor volumen de información resultaron más problemáticas. Los casos las instalaciones *Ameghiniana* (65 GB de datos) y Portal de Revistas de la UNLP (60 GB), son los más representativos aquí.
- 2. Antigüedad de la instalación:** el uso de versiones muy antiguas de PHP y MySQL requiere dedicar un esfuerzo adicional, ya que surgen problemas vinculados a sistemas de codificación (UTF8, Latin 8859, etc.) y a librerías particulares (por ejemplo: el uso de PDO o MySQLi como librería de conexión entre PHP y MySQL). Todas las actualizaciones desde OJS 2.x hacia OJS 3.X presentaron problemas.
- 3. Historial de actualizaciones:** las instalaciones que ya superaron varios procesos de actualización a lo largo del tiempo pueden presentar problemas de inconsistencia en los datos (por ejemplo: una revisión sin su revisor, una descarga de archivo sin el correspondiente archivo, etc.). Este tipo de problemas se presentó principalmente en la instalación del Portal de Revistas de la UNLP, que desde su lanzamiento en el año 2008 ya superó muchas actualizaciones (VILLARREAL, 2013, 2014, 2017; PERCIVALE, 2015).

## Conclusiones

A lo largo de este trabajo presentamos una herramienta que permite automatizar algunas de las tareas requeridas para realizar procesos de actualización del software OJS. El uso de esta herramienta reduce la complejidad del proceso iterativo de actualización y migración; destacamos la incorporación de las tecnologías de código abierto Docker, Docker-compose,

GNU Make y Bash, como herramientas para abstraer y simplificar la cantidad de instrucciones que requiere este proceso, y de emular los distintos entornos de ejecución requeridos para las diferentes versiones de OJS. Si bien esta propuesta no automatiza por completo la actualización de OJS, ya que se encontrarán errores particulares de cada instalación de OJS que deben ser revisados, para luego generar las consultas necesarias que los resuelvan, se reducen notablemente las tareas de reproducción de los procesos de actualización y se minimizan los errores introducidos al evitar la escritura de comandos complejos en cada etapa del proceso.

Asimismo, resulta evidente que es posible llevar esta herramienta a un mayor grado de automatización. Como futuras líneas de investigación y desarrollo, sería muy interesante incorporar a SUM OJS la posibilidad de ejecutar actualizaciones entre versiones muy alejadas de manera desatendida, es decir, poder actualizar por ejemplo de la versión 2.4 a la versión 3.3 con un solo comando, que internamente ejecutará cada uno de los comandos propuestos aquí para restaurar la base de datos, ejecutar la migración, generar backups y volver a repetir con la siguiente versión de OJS. Es claro que esto requerirá un mayor esfuerzo de configuración por parte del usuario, ya que se deberá generar un entorno virtual para cada versión de OJS a la que se intentará migrar, pero está claro que una vez generados dichos entornos, el proceso podría ejecutarse de manera íntegra en un sólo comando.

Todas las herramientas descritas en este documento están disponibles bajo licencias de código abierto. El código fuente del proyecto se encuentra accesible desde el repositorio GitHub: <https://github.com/sedici/script-ojs>

## Bibliografía

- BAKER, P. (2020). Using GNU Make to Manage the Workflow of Data Analysis Projects. *Journal of Statistical Software*, 94, 1-46. <https://doi.org/10.18637/jss.v094.c01>
- BROBERG, M. (2019) The birth of the Bash shell. OpenSource.com <https://opensource.com/19/9/command-line-heroes-bash>
- DI TOMMASO, P., PALUMBO, E., CHATZOU, M., PRIETO, P., HEUER, M. L., & NOTREDAME, C. (2015). The impact of Docker containers on the performance of genomic pipelines. *PeerJ*, 3, e1273. <https://doi.org/10.7717/peerj.1273>
- DOWNLOAD | PUBLIC KNOWLEDGE PROJECT. [https://pkp.sfu.ca/ojs/ojs\\_download/](https://pkp.sfu.ca/ojs/ojs_download/)
- GNU MAKE. <https://www.gnu.org/software/make/manual/make.html>
- IBRAHIM, M. H., SAYAGH, M. & HASSAN, A. E. (2021). A study of how Docker Compose is used to compose multi-component systems. *Empir Software Eng* 26, 128. <https://doi.org/10.1007/s10664-021-10025-1>
- PERCIVALE, B. (2015). Actualización del Portal de Revistas de la UNLP: optimizado para móviles. [Blog] 6 de noviembre de 2015 <https://blog.sedici.unlp.edu.ar/2015/11/06/actualizacion-del-portal-de-revistas-de-la-unlp-optimizada-para-moviles/>
- TAGS. PKP/OJS <https://github.com/pkp/ojs/tags>
- VILLARREAL, G. L. (2013). Actualización del Portal de Revistas de la UNLP. [Blog] 8 de marzo de 2013. <https://blog.sedici.unlp.edu.ar/2013/03/08/actualizacion-del-portal-de-revistas-de-la-unlp/>
- VILLARREAL, G. L. (2014). Nuevas características de OJS 2.4.5. [Blog] 28 de octubre de 2014. <https://blog.sedici.unlp.edu.ar/2014/10/28/nuevas-caracteristicas-de-ojs-2-4-5/>
- VILLARREAL, G. L. (2017). Actualización a OJS 3 del Portal de Revistas de la UNLP. [Blog] 7 de febrero de 2017. <https://blog.sedici.unlp.edu.ar/2017/02/07/actualizacion-a-ojs-3-del-portal-de-revistas-de-la-unlp/>

# Es necesario un identificador persistente además del DOI (Digital Object Identifier) para el sur global

## CARLOS NORBERTO AUTHIER

Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) (Argentina)

[cauthier@conicet.gov.ar](mailto:cauthier@conicet.gov.ar)

## DIEGO FERREYRA

Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) (Argentina)

[dferreyra@conicet.gov.ar](mailto:dferreyra@conicet.gov.ar)

### RESUMEN

Aborda la necesidad de otros identificadores persistentes para los artículos científicos en el sur global. Es posible afirmar en la actualidad que el DOI (Digital Object Identifier) se ha transformado en un estándar de alcance mundial para la identificación de los artículos científicos. Esta presentación muestra una experiencia desarrollada en Argentina con soporte del Consejo Nacional de Investigaciones Científicas y Técnicas, la cual está basada en un identificador persistente que no tiene costo para los editores científicos. El servicio está basado en el ARK (Archival Resource Key), desarrollado por John Kunze de la California Digital Library (<https://arks.org/>). Al 2022, un total de 67 revistas argentinas cuentan con este identificador con las mismas características y funciones operativas del DOI. La necesidad de desarrollar identificadores persistentes sin costo para los editores es aún más pertinente en el modelo de Acceso Abierto que no cobra APC (Article Processing Charge) a los autores.

## **PALABRAS CLAVE**

Identificadores persistentes, PID, DOI, ARK.

*Persistent identifiers, PID, DOI, ARK.*

## **Acerca de ARK-CAICYT**

Es abundante la bibliografía sobre los identificadores persistentes y aún más sobre el estándar actual, el Digital Object Identifier (DOI) y su importancia para la identificación de objetos digitales. Como principal fuente de documentación es posible recurrir al DOI handbook, que está disponible en la página de la organización (INTERNATIONAL DOI FOUNDATION, 2013). También hay bibliografía disponible del identificador Handle.Net (HNR) en su página oficial, en este caso un manual técnico (HANDLE, 2018). Además, entre los cuatro identificadores persistentes reconocidos por la International Federation of Library Associations and Institutions (IFLA), debemos mencionar el Persistent Uniform Resource Locator (PURL), que fue desarrollado e implementado por el Online Computer Library Center (OCLC).

Es posible consultar la información sobre los identificadores persistentes en el artículo de ZUMER (2008), donde se describen los principios que los rigen, los requerimientos, la funcionalidad, etc.

En el caso del identificador elegido para esta iniciativa, el Archival Resource Key (ARK), que es uno de los cuatro reconocidos por IFLA, es posible recurrir al desarrollador JOHN KUNZE (2003), quien explica las ventajas de este identificador. Este puede adjudicarse no solo a artículos científicos sino también a imágenes, textos, conjuntos de datos y herramientas de búsqueda. Ha sido implementado por la Biblioteca Digital de California (CDL) y al 2018 por 550 organizaciones y bibliotecas entre las que se cuenta la Bibliothèque Nationale de France.

## ¿Qué es el ARK?

¿Es posible, siendo el estándar para artículos científicos el identificador DOI implementar en Latinoamérica un identificador alternativo gratuito? De la misma manera que Latinoamérica ha sido pionera en la implementación del Acceso Abierto (en inglés, Open Access u OA), nos propusimos ante la demanda de los editores científicos en la Argentina que tenían dificultades para que sus instituciones abonaran el costo del DOI (U\$S 275 anuales más el costo por artículo), desde el Centro Argentino de Información Científica y Tecnológica (CAICYT), dependiente del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) de la República Argentina, una experiencia piloto para asignar un identificador único y persistente a los artículos científicos.

## ¿Por qué usar ARK en vez de DOI?

En primer lugar, las revistas científicas de acceso diamante suelen tener problemas para conseguir fondos, por lo tanto la ventaja del ARK es que no tiene el costo del DOI. Entre otras ventajas del ARK, es posible mencionar que se pueden crear identificadores sin metadatos, es posible crear un identificador incluso antes que exista un objeto digital, también es posible mantener el identificador privado mientras los datos y metadatos evolucionan. Luego puede asignarse un identificador adicional, como un DOI, un handle, etc. Además, el protocolo ARK permite crear identificadores más cortos, ya que las mayúsculas y minúsculas permiten cadenas más densas. Y, fundamentalmente, utilizar un identificador persistente con estructura abierta coherente con los valores de la mayoría de las organizaciones científicas latinoamericanas.

## Relevamiento de la utilización de identificadores persistentes

Recientemente, se realizó un relevamiento de la utilización de identificadores persistentes basado en Directory of Open Access Journals (DOAJ) al 11 de mayo de 2022. En él es posible obtener los datos públicos de las revistas científicas en formato JSON (Java Script Object Notation), que son volcados semanalmente. El directorio incluye los datos de 130 países en 80 idiomas y al día indicado releva los datos de 17.602 revistas. DOAJ es un directorio en línea que brinda acceso a revistas revisadas por pares de alta calidad y acceso abierto. Todos los datos están disponibles gratuitamente. De los datos obtenidos es posible comprobar que un 63 % cuentan con el DOI como identificador persistente. Un número menor, no superior en ningún caso al 1 % cuenta con los identificadores ARK, Handle y PURL. Y 6.185 revistas, es decir un 35 % no cuentan con ningún identificador para artículos científicos.

La elección de ese directorio está basada en su cobertura mundial y el prestigio, y la posibilidad de descargar los datos completos de las revistas de manera gratuita. Por otra parte, los datos muestran la marginalidad del uso de otros identificadores. El más utilizado en segundo lugar es el handle (<https://www.handle.net/>). Tiene un costo menor que el DOI pero solo 49 revistas científicas lo utilizan, lo que representa un 0,27 % del total de revistas. En tercer lugar, se ubican las revistas que utilizan el ARK, identificador que, como se dijo, no tiene costo para las revistas, pero ellas deben instrumentar su propio resolvidor de identificadores. Solo 16 revistas científicas utilizan este identificador, en porcentaje un escaso 0,09 %. En cuarto lugar, cuatro revistas científicas utilizan el PURL, lo que representa el 0,03 % del total.

## Identificadores persistentes

Los identificadores persistentes constituyen una infraestructura primaria para la representación, formalización, circulación y operacionalización del

conocimiento científico, ofreciendo mecanismos para la identificación no ambigua, persistente y funcional de los componentes intervinientes en las prácticas de investigación, desarrollo y comunicación científica. Permiten referenciar de manera formal y unívoca artefactos de todo tipo, ya sean constructos abstractos, entidades físicas, personas, instituciones o componentes de mediaciones comunicacionales.

Establecen las condiciones de posibilidad para que un componente del sistema científico pueda ser identificado, representado y utilizado, facilitando el reúso, citación y socialización de producciones, herramientas y resultados.

Catálogos, nomencladores, convenciones y estándares son ejemplos de las estrategias que a lo largo del tiempo desarrollaron las comunidades en su esfuerzo por designar de manera sistemática y no ambigua el conocimiento construido. El desafío que enfrentamos actualmente nos invita a diseñar y adoptar un esquema de identificadores persistentes en un medio digital y altamente integrado a través de redes informáticas globales. En tal sentido, una estrategia coherente con el actual contexto socio-técnico requiere un modelo que reúna las siguientes condiciones:

- **Referencia no ambigua en contextos digitales en red:** debe poder referenciar una entidad digital de manera no ambigua en el contexto global de Internet.
- **Resolución funcional:** debe poder asociarse con un mecanismo de resolución que garantice la disponibilidad y acceso hacia las entidades digitales referenciadas.
- **Persistencia:** la relación entre la referencia y la entidad digital referenciada debe mantenerse a través del tiempo.

Existen diferentes iniciativas y proyectos en desarrollo en la actualidad orientados a consolidar estándares y mecanismos capaces de cumplir con las condiciones antes mencionadas. Con el fin de establecer un instrumento de relevamiento, hemos establecido una serie de criterios de análisis y evaluación compatibles con una agenda de política científica situada en nuestras realidades y articular con modelos de ciencia abierta.

Los criterios de análisis a considerar son:

- **Sustentable:** capacidad para resultar sostenible desde el punto de vista organizativo y económico.
- **Técnicamente interoperable:** disponibilidad de un diseño tecnológico articular con las infraestructuras de red disponibles, los mecanismos de provisión de servicios automatizados y los estándares digitales de descripción de recursos
- **Tecnológicamente neutro:** condiciones para operar de manera no-dependiente con respecto a tecnologías de base específicas u otras limitaciones (datos, metadatos o servicios propietarios o restrictivos).
- **Políticamente fiable:** capacidad para evitar obstáculos legales (restricciones de acceso o bloqueo de servicios).
- **Portabilidad funcional:** condiciones para su utilización en diversos contextos y medios.
- **Alcance:** universo de entidades que es capaz de referenciar (objetos digitales, físicos, partes de objetos, etc.).
- **Documentación:** es un estándar documentado o un estándar de facto.
- **Soberanía:** modelo de gobierno que lo regula y administra.
- **Robustez:** arquitectura de persistencia utilizada.
- **Interoperabilidad diacrónica:** condiciones de articulación con infraestructuras previas o futuras.

## El rol del CAICYT y Centro ISSN

En el CAICYT además funciona el Centro Nacional de ISSN, que otorga los números de identificación unívoca para las publicaciones seriadas argentinas. Por lo que en este desarrollo, elegimos vincular el estándar ARK con el estándar ISSN.

## Cómo funciona

1. La revista científica argentina se pone en contacto con el CAICYT y este le asigna una cuenta de usuario con una contraseña con la cual podrá ingresar a través de la siguiente URL:  
<http://id.caicyt.gov.ar/issn/>
2. Allí se le solicitarán los datos de cuenta de usuario para ingresar (correo electrónico, contraseña).
3. Una vez validado el usuario, accederá a la pantalla de ISSN asignados a dicha cuenta, donde se expondrá la lista de ISSN y sus nombres de revista.
4. Al hacer clic sobre el nombre de alguna de las revistas de la lista, se accederá a la pantalla de Lista de recursos asociados a dicho ISSN, y Formulario para nuevo recurso.
5. Esta pantalla dispone de dos secciones: **a.** Formulario para agregar un nuevo recurso/artículo asociándolo al ISSN actual. **b.** Lista de Recursos ya asignados a dicho ISSN.
  - 5a. En el campo "URL del recurso" se debe ingresar la URL del artículo de la revista sin ninguna alteración y luego hacer click en el botón "Generar ARK".
  - 5b. Luego de unos instantes la página se actualizará mostrando justo debajo de la URL ingresada el recientemente generado identificador ARK-CAICYT para dicha URL.

## Requisitos de postulación

- La revista debe tener carácter científico o técnico.
- La publicación debe tener un ISSN de Argentina.
- Contar con un ISSN para la versión en línea (en caso de no tenerlo puede tramitarlo en el CAICYT).
- Estar en línea al momento de solicitar su inclusión en el proyecto en su URL definitiva.

- Tener dos años de antigüedad en su versión en línea.
- Contar con un sitio web oficial o institucional dedicado.
- Adherir a los principios de Acceso Abierto ([http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/publications/policy\\_guidelines\\_oa\\_sp\\_reduced.pdf](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/publications/policy_guidelines_oa_sp_reduced.pdf))
- No se incluirán revistas digitalizadas en un solo archivo PDF, deben mínimamente tener una tabla de contenidos y los artículos separados en PDF individuales.
- Los responsables editoriales y miembros de los comités deben proporcionar un correo electrónico.
- Sistema de arbitraje externo a la entidad editora.

Los datos requeridos para participar son:

1. ISSN de la revista
2. Nombre de contacto
3. Correo electrónico del contacto
4. CUIL del contacto

Publicaciones/usuarios del servicio de Identificadores persistentes ARK-CAICYT:

- *I+A Investigación + Acción*
- *Páginas de Filosofía*
- *Revista Argentina de Coloproctología*
- *Registros: Revista de Investigación Histórica*
- *Otra Economía*
- *Revista Latinoamericana de Estudios Rurales*
- *Diccionario de Científicos Argentinos "Dra. Cecilia Grierson"*
- *ARTILUGIO*
- *Revista TEFROS*
- *Cuadernos de Economía Crítica*
- *RASAL Lingüística*

- *Toma Uno*
- *AVATARES de la Comunicación y la Cultura*
- *Sudamérica: Revista de Ciencias Sociales*
- *Anuario Centro de Estudios Económicos de la Empresa y el Desarrollo*
- *Costos y Gestión*
- *Cuadernos Universitarios*
- *Cuadernos de Ingeniería*
- *OMNIA: Derecho y sociedad*
- *Teks del Sud: Cuadernos de Arquitectura y Diseño*
- *Crítica y Resistencias*
- *Mundo de Antes*
- *Revista Ucronías*
- *Argumentos: Revista de Crítica Social*
- *Revista de la Sociedad Odontológica de La Plata*
- *Ab Intus*
- *Contextos de Educación*
- *Cronía*
- *Temas y Problemas de Comunicación*
- *Cultura en Red*
- *Revista Argentina de Medicina (En línea)*
- *El Cardo (Paraná. En línea)*
- *Del Prudente Saber y el Máximo Posible de Sabor (En línea)*
- *Revista Latinoamericana de Antropología del Trabajo*
- *Sociedad y Religión*
- *(COORD) Coordinadas. Revista de Historia Local y Regional*
- *Revista del Instituto Internacional de Costos*
- *Educación y Vínculos: Revista de Estudios Interdisciplinarios en Educación*
- *Horticultura Argentina*
- *Chuy. Revista de Estudios Literarios Latinoamericanos*
- *Quintú Quimün. Revista de Lingüística*
- *(En)clave Comahue. Revista Patagónica de Estudios Sociales*
- *Anuario Pilquen. Sección Divulgación Científica*
- *Revista Pilquen. Sección Psicopedagogía*
- *Revista Pilquen. Sección Ciencias Sociales*
- *RAC - Revista Argentina de Comunicación*

- *Revista de la Escuela de Perfeccionamiento en Investigación Operativa*
- *Campo Universitario: Revista de Educación Superior*
- *Trabajo y Sociedad*
- *Revista Hormigón*
- *Revista Latinoamericana de Economía y Sociedad Digital*
- *Revista Enfermería Neonatal*
- *Revista Electrónica de Fuentes y Archivos (REFA)*
- *Revista de Ciencias Empresariales y Sociales*
- *Revista Científica de UCES (RCU)*
- *Desvalimiento Psicosocial*
- *Ratio Iuris. Revista de Derecho Privado*
- *Subjetividad y Procesos Cognitivos*
- *Conflicto Social*
- *Revista Educación en Biología*
- *Vivomatografías*
- *La Zaranda de Ideas*
- *Revista de Arqueología Histórica Argentina y Latinoamericana (RDAHAYL)*

## Costo y arquitectura de un resoledor nacional

El costo del desarrollo es, aproximadamente, el que le demanda a un desarrollador senior 90 días de trabajo full time. Estos tiempos varían en el caso de un desarrollador semi senior (150 días) o junior (200 días). Las arquitecturas utilizadas fueron HTML en conjunción con CSS para el desarrollo frontend y JavaScript para el desarrollo del backend. Para minimizar los riesgos de seguridad se desarrollarán algoritmos de encriptación y servidores espejados para mantener las bases de datos.

Además, este desarrollo técnico fue acompañado de capacitaciones y asesoramiento a los usuarios. Actualmente se dispone de un tutorial disponible en línea ([http://id.caicyt.gov.ar/docs/ARK\\_tutorial.pdf](http://id.caicyt.gov.ar/docs/ARK_tutorial.pdf)).

Para acompañar las capacitaciones se desarrollaron sucesivos materiales disponibles en línea, en el 2018: ARK como identificador persistente gratuito: una propuesta. En el 2022: Identificadores persistentes para la ciencia argentina: ARK-CAICYT.

## Historia del ARK-CAICYT y futuros pasos

Durante el 2018 se realizó el diseño técnico y el desarrollo informático. En el 2019 se publicó el prototipo y se desarrolló el pilotaje automático y finalmente la implementación de la plataforma de gestión remota. En el 2022 se implementó de manera pública el backend de gestión de usuarios. Está previsto la implementación de API y el esquema de redundancia para el año 2023.

## Conclusión

La Argentina ha demostrado que se puede implementar un identificador persistente sin costo para la ciencia con un proyecto que ya cuenta con más de sesenta revistas científicas, con todos los requisitos técnicos de los estándares internacionales. Es revelador el texto publicado por John Kunze en el blog de la página oficial del ARK en 2021, donde argumenta sobre los diez mitos persistentes sobre los identificadores persistentes (KUNZE, 2021), en el mito 9 argumenta: los PID deben estar centralizados. Falso. Cualquier PID con un núcleo único globalmente después del nombre de host de la URL es persistente. De hecho, si no puede ser atendido por otros hosts, no puede persistir ya que su destino está ligado a un host. Ningún nombre de host o protocolo dura para siempre. El ARK-CAICYT ha sido un gran impulso para la Ciencia Abierta en la Argentina que no podía costear el DOI. Puede replicarse en otros países del sur global con muy poca inversión de parte de los Estados. Estamos a disposición de ellos para compartir la iniciativa.

## Bibliografía

- AUTHIER, CARLOS NORBERTO, FERREYRA, DIEGO ANDRÉS, Y BIGLIERI, HERNÁN. (2018). ARK como identificador persistente gratuito: una propuesta. COMCIENT: repositorio institucional del CAICYT especializado en información y comunicación científica. <http://www.caicyt-conicet.gov.ar/comcient/ark%3A/16680081/rscggh>
- BERMÈS, E. (2006). Des identifiants pérennes pour les ressources numériques: l'expérience de la BnF. <https://www.ifla.org/files/assets/pac/ipn/ipnn40.pdf>
- DIRECTORY OF OPEN ACCESS JOURNALS (DOAJ). <https://doaj.org/docs/public-data-dump/>
- DOI FOUNDATION. (2013). DOI (Digital Object Identifier): The Foundation. <http://www.doi.org/>
- FERREYRA, DIEGO ANDRÉS Y AUTHIER, CARLOS NORBERTO. (2022). Identificadores persistentes para la ciencia argentina: ARK-CAICYT. COMCIENT: repositorio institucional del CAICYT especializado en información y comunicación científica. <http://www.caicyt-conicet.gov.ar/comcient/ark%3A/16680081/rscggr>
- HANDLE (2018). Technical Manual Version 9 Preliminary edition. [http://www.handle.net/tech\\_manual/HN\\_Tech\\_Manual\\_9.pdf](http://www.handle.net/tech_manual/HN_Tech_Manual_9.pdf)
- KUNZE, J. (2021). Ten persistent myths about persistent identifiers. <https://arks.org/blog/ten-persistent-myths-about-persistent-identifiers/>
- KUNZE, J. (2003). Towards electronic persistence using ARK identifiers. In: Proceedings of the 3rd ECDL Workshop on Web Archives.
- ZUMER, M. (2008). Guidelines for National Bibliographies in the Electronic Age. <https://www.ifap.ru/pr/2008/n080616b.pdf>

# Evaluación de un repositorio institucional a través de NDSA Levels: Caso CIC Digital

## **SANTIAGO TETTAMANTI**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[santit@sedici.unlp.edu.ar](mailto:santit@sedici.unlp.edu.ar)

## **MARISA RAQUEL DE GIUSTI**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[marisa.degiusti@sedici.unlp.edu.ar](mailto:marisa.degiusti@sedici.unlp.edu.ar)

## **ARIEL JORGE LIRA**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[ariel.lira@sedici.unlp.edu.ar](mailto:ariel.lira@sedici.unlp.edu.ar)

## **RESUMEN**

Este trabajo está dedicado a relatar la experiencia de autoevaluación de un repositorio institucional realizada utilizando NDSA. Se inicia a partir de la exposición del contexto de los repositorios institucionales y los estándares utilizados para auditar su confiabilidad, en cuanto al mantenimiento del acceso y comprensión de sus contenidos a largo plazo. La propuesta de niveles de NDSA abarca diversas áreas del repositorio, es muy comprensible y brinda un primer acercamiento al tema de preservación más sencillo que otros estándares como la ISO 16363. El trabajo construye la matriz de resultados del test del repositorio CIC Digital, así como una propuesta de acciones para poder mejorar los aspectos no cumplidos o parcialmente cumplidos.

## **PALABRAS CLAVE**

Repositorio institucional; preservación digital; estándares de evaluación; autoevaluación.

*Institutional repository; digital preservation; evaluation standards; self-evaluation.*

## **Introducción**

### **Contexto y antecedentes**

Los repositorios institucionales de acceso abierto se han convertido en herramientas esenciales para la comunicación académica en la era digital; además de servir como una colección digital que agrupa la producción académica y científica de una institución, ofrecen servicios sobre esta producción y aseguran el acceso abierto y a largo plazo de sus obras. Para garantizar dicho acceso, la preservación digital juega un papel fundamental. Según FERRERAS-FERNÁNDEZ (2010), la preservación digital se refiere a la conservación para asegurar la accesibilidad, la recuperación y el uso a futuro de los materiales digitales, y a las distintas técnicas y actividades dirigidas a alcanzar esos objetivos. Según la autora, para que un repositorio logre la perdurabilidad de sus objetos, no sólo se necesitan esfuerzos técnicos sino que existen otros retos: legales (permiso de los autores), económicos e institucionales.

TÉRMENS y LEIJA (2017) explican que una parte vital del proceso de preservación digital en un repositorio es poder valorar si los sistemas, actividades y técnicas de preservación que se utilizan son los correctos; para ello es necesario utilizar distintas herramientas y metodologías que sirven para evaluar el grado de cumplimiento con estándares y buenas prácticas existentes. Los autores señalan, además, a los sistemas de auditoría como una de las metodologías más extendidas para conseguir repositorios confiables;

éstos permiten determinar si un repositorio de preservación es seguro, y por tanto, si es confiable, pero están diseñados para evaluar sistemas de preservación existentes y no para la detección de acciones a implementar para lograr la preservación, objetivo común a muchas organizaciones pequeñas. Sumado a esto, el uso de auditorías tradicionales suele ser caro y de una complejidad alta que requiere personal experto para llevarlas a cabo. En ese sentido, las organizaciones pequeñas quizás prefieran otro tipo de sistemas de evaluación, con mayor facilidad de aplicación y con métodos que permitan la autoevaluación.

### **Sobre NDSA Levels**

La National Digital Stewardship Alliance (NDSA) fue creada en 2010 como una iniciativa de la Library of Congress. Se trata de un consorcio de organizaciones comprometidas con la conservación a largo plazo de la información digital. A la fecha cuenta con 267 instituciones miembro, entre las que se encuentran universidades, organizaciones gubernamentales, organizaciones sin fines de lucro, empresas comerciales y asociaciones profesionales.

En 2013, la NDSA creó Levels of Digital Preservation (en 2019 fue lanzada la versión 2.0), un recurso para ayudar a aquellas instituciones o personas que deseen evaluar su sistema de preservación digital. NDSA Levels se compone de una matriz sobre la cual se definen tareas, actividades o normas que el repositorio o sistema debería llevar a cabo para asegurar la preservación y el acceso a largo plazo. Estas tareas se dividen en cinco categorías distintas:

- Almacenamiento y localización geográfica
- No alteración de ficheros/archivos e integridad de los datos
- Seguridad de la información
- Metadatos
- Formatos de ficheros/archivos

Por cada categoría se definen a su vez cinco niveles de cumplimiento de las tareas. La estructura general de la matriz es progresiva: las acciones en el primer nivel son requisitos previos necesarios para aquellos en los niveles de segundo a cuarto o son en sí mismas las actividades más urgentes a lograr. En términos generales, a medida que se avanza en cada uno de los niveles, del Nivel 1 al Nivel 4, se está moviendo de la necesidad básica de garantizar la conservación de bits, a requisitos más amplios para realizar un seguimiento del contenido digital y poder garantizar que pueda estar disponible a través de períodos más largos de tiempo. En el ANEXO 1 se deja la traducción al español de la matriz de niveles de preservación digital.

Se debe tener en cuenta que la propuesta de NDSA se centra en la posibilidad de una autoevaluación por parte de los repositorios, ofrece una visión simplificada de las tareas de preservación y no un listado exhaustivo como, por ejemplo, se puede encontrar en Nestor, TRAC o ISO 16363. En consecuencia, las respuestas pueden ofrecer una visión más optimista de la que correspondería a la situación real, ya que se ignoran otras facetas claves como la robustez institucional o la sostenibilidad económica (TÉRMENS & LEIJA, 2017) que sí cubre por ejemplo ISO 16363.

La NDSA provee, además de traducciones a distintos idiomas de la matriz, una guía de evaluación en la que documenta cómo se debería evaluar un sistema con su uso, junto con una plantilla en formato de hoja de cálculo (NDSA LEVELS OF PRESERVATION ASSESSMENT SUBGROUP, 2019), a modo de ayuda para completar en qué medida nuestro sistema cumple con las tareas o requerimientos presentes en la matriz.

## **Evaluación de CIC Digital**

### **Metodología**

Se realizó una evaluación a partir de la matriz de preservación de NDSA del repositorio institucional CIC Digital, de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires. Si bien se sugiere responder de

manera afirmativa o negativa a las recomendaciones que se encuentran en la matriz, puede suceder que para alguno de los puntos planteados solo se haya completado una parte, o si el punto está dividido en subtarefas, se hayan completado alguna de ellas, es decir, no se satisface el punto en su totalidad pero, a la vez ya se ha implementado alguna parte de lo recomendado. Para tener en cuenta dichos casos, se decidió seguir la metodología propuesta por TÉRMENS y LEJÍA, la cual consiste en marcar con un color especial las respuestas afirmativas en la tabla, pero con el agregado de marcar con otro color las respuestas en las que se ha completado el punto de manera parcial, y con otro las respuestas negativas. Las respuestas que se marcaron como completadas de manera parcial no sumaron puntos en el recuento final de resultados, solo sumaron puntos las respuestas afirmativas. De esa manera se puede apreciar rápidamente y a golpe de vista los resultados de manera general.

La evaluación de cada uno de los puntos planteados y el resultado final de cada uno fue tomado de manera conjunta por varios integrantes de los distintos equipos que trabajan en CIC Digital. Se realizaron sucesivas reuniones con al menos una persona encargada de alguna de las actividades involucradas en la preservación digital en el repositorio, o al menos que sean referenciadas por la matriz de evaluación. Entre los equipos se pueden mencionar al equipo de infraestructura del repositorio, responsable entre otras cosas de las copias de seguridad, chequeos de integridad, del mantenimiento y seguridad de los distintos servidores; el equipo de desarrollo, encargado de los chequeos por software (por ej, virus, integridad y formato de archivos), de la seguridad del sistema, de la base de datos, de la curación automática de los metadatos, etc.; el equipo administrativo, que se ocupa de controlar el contenido que ingresa al repositorio, los tipos y cantidad de metadatos usados, la variedad de formatos; y los encargados de la gestión del repositorio, responsables de la visión general de este, las políticas presupuestarias y las decisiones institucionales. Puede darse el caso, en especial en repositorios pequeños que no poseen gran cantidad de personal disponible, que una misma persona realice actividades correspondientes a varios de estos equipos.

En este trabajo, se debatió de manera conjunta acerca del cumplimiento o no de las recomendaciones de la matriz, se realizó una evaluación de cada punto por separado y luego, en caso de no cumplimiento del requerimiento planteado, se fijaron las acciones a llevar adelante. Para algunas recomendaciones, por lo general las correspondientes al nivel 4 de la matriz, no fue posible plantear una solución viable para el repositorio, muchas veces esto ocurre por falta de recursos de infraestructura tecnológica o de presupuesto.

### **Matriz de evaluación NDSA Levels para CIC Digital**

Aquí se presenta el análisis de preservación digital sobre el repositorio CIC Digital, realizado con la matriz de evaluación de NDSA. El color verde significa que la recomendación está satisfecha en su totalidad, el amarillo indica que gran parte del punto planteado fue resuelto, pero que hay alguna parte de la recomendación que no fue satisfecha en su totalidad, y el color rojo muestra aquellos puntos que no fueron resueltos aún por el repositorio.

	NIVEL 1 (PROTEJA SUS DATOS)	NIVEL 2 (CONOZCA SUS DATOS)	NIVEL 3 (CONTROLE SUS DATOS)	NIVEL 4 (REPARE SUS DATOS)	PUNTAJE
<b>Almacenamiento</b>	Tener dos copias completas en ubicaciones separadas	Tener tres copias completas con al menos una copia en una ubicación geográfica distinta	Tener al menos una copia en una ubicación geográfica con amenaza de desastre diferente a las otras copias	Tener al menos tres copias en ubicaciones geográficas distintas, cada una con una amenaza de desastre diferente.	<b>0/4</b>
	Documentar todos los medios de almacenamiento donde esté almacenado el contenido	Documentar el almacenamiento y medios de almacenamiento, indicando los recursos y las dependencias que estos requieren para funcionar	Rastrear la obsolescencia del almacenamiento y los medios	Maximizar la diversificación del almacenamiento para evitar puntos únicos de falla	
	Poner el contenido en soportes de almacenamiento estables		Tener al menos una copia en un medio de almacenamiento de diferente tipo	Tener un plan y realizar acciones para abordar la obsolescencia del hardware, software y medios de almacenamiento	
<b>No alteración de archivos e integridad de los datos</b>	Verificar que la información de integridad se ha proporcionado con el contenido	Verificar la información de integridad al mover o copiar contenido	Verificar la información de integridad del contenido en intervalos fijos	Comprobar la integridad de todo el contenido en respuesta a situaciones o actividades específicas.	<b>1/4</b>
	Generar información de integridad si esta no ha sido proporcionada con el contenido	Usar bloqueadores de escritura cuando se trabaja con medios originales	Documentar los procesos y resultados de verificación de información de integridad	Verificar la información de integridad en respuesta a eventos o actividades específicas	
	Se verifica virus en todo el contenido; se aísla el contenido en cuarentena según sea necesario	Hacer una copia de seguridad de la información de integridad y almacenar una copia en una ubicación separada del contenido	Realizar una auditoría de la información de integridad bajo demanda	Reemplazar o reparar el contenido dañado según sea necesario	

	<b>NIVEL 1 (PROTEJA SUS DATOS)</b>	<b>NIVEL 2 (CONOZCA SUS DATOS)</b>	<b>NIVEL 3 (CONTROLE SUS DATOS)</b>	<b>NIVEL 4 (REPARE SUS DATOS)</b>	<b>PUNTAJE</b>
<b>Seguridad de la información</b>	Se determinan los agentes humanos y de software que deben estar autorizados para leer, escribir, mover y eliminar contenido	Documentar los agentes humanos y de software autorizados para leer, escribir, mover y eliminar contenido y aplicar estos cambios	Mantener los registros (logs) e identificar a los agentes humanos y de software que realizaron acciones sobre el contenido	Se realizan revisiones periódicas de acciones/registros (logs) de acceso	<b>1/4</b>
<b>Metadatos</b>	<p>Crear un inventario de contenido, documentando también la ubicación de almacenamiento actual de estos</p> <p>Hacer una copia de respaldo del inventario y almacenar al menos una copia por separado</p>	Almacenar suficientes metadatos para saber cuál es el contenido (esto podría incluir alguna combinación de aspectos administrativos, técnicos, descriptivos, de preservación y estructurales)	<p>Determinar qué estándares de metadatos aplicar</p> <p>Encontrar y completar los vacíos en sus metadatos para cumplir con esos estándares</p>	Registrar las acciones de preservación asociadas con el contenido y cuando ocurren esas acciones Implementa los estándares de metadatos elegidos	<b>3/4</b>
<b>Formatos de archivos</b>	Documentar los formatos de archivo y otras características de contenido esenciales, incluidos cómo y cuándo fueron identificados	<p>Verificar los formatos de archivo y otras características de contenido esenciales</p> <p>Establecer relaciones con los creadores de contenido para fomentar la elección sostenible de archivos</p>	Monitorear la obsolescencia y los cambios en las tecnologías de las que depende el contenido	Realizar migraciones, normalizaciones, emulación y actividades similares que garanticen el acceso al contenido.	<b>4/4</b>
<b>Puntaje global</b>	<b>3/5</b>	<b>2/5</b>	<b>3/5</b>	<b>1/5</b>	<b>9/20</b>

## Descripción del cuadro

Se puede observar, en base a los resultados, que CIC Digital cumple casi con la mitad de las recomendaciones de la matriz, 9 de las 20 recomendaciones fueron marcadas como completadas, es decir, el 45 % de la matriz. El nivel 1 se satisface en todos sus puntos excepto en dos; los niveles 2 y 3 tienen tres y cuatro puntos no satisfechos respectivamente, aunque por como se da su distribución, el nivel 2 cumple solo con la totalidad de los puntos en dos categorías mientras que el nivel 3 lo consigue en tres; el nivel 4 es el que tiene menos recomendaciones satisfechas, con solo la categoría "Formato de archivos", en donde todos los puntos se cumplen, esta baja performance en el último nivel es esperable si se tiene en cuenta que la estructura general de la matriz es progresiva y que por lo general las actividades de los primeros niveles son necesarias para cumplir con los niveles subsiguientes.

En lo que respecta a la categoría de "Almacenamiento", CIC Digital posee copias de seguridad tanto en servidores propios, como en servidores que se encuentran en otra dependencia de la UNLP. Esta dependencia se encuentra en una ubicación geográfica distinta a la de los servidores del repositorio, y los servidores se encuentran por encima del primer piso de sus respectivos edificios (esto evitaría riesgos como la inundación), lo que permite cumplir tres de las recomendaciones propuestas para esta categoría; sin embargo, la distancia que separa las dos localizaciones con copias de seguridad no es lo suficientemente grande como para evitar que ambas sean afectadas por un desastre natural que involucre a toda la ciudad en donde se encuentran.

En lo que refiere al control de la obsolescencia de los medios de almacenamiento, la renovación y mantenimiento del hardware se realiza pero de manera manual y esporádica, pues no hay un plan de renovación o de control de renovación regular del equipamiento, ni tampoco existe una planificación presupuestaria que lo contemple. Algo común a los puntos de esta categoría es la falta de documentación formal, no se cuenta con un plan de riesgos ni tampoco se encuentran documentados los distintos medios y

equipos en los que se almacena el contenido y con los que se tiene en funcionamiento el repositorio.

La segunda categoría, “No alteración de archivos e integridad de los datos”, solo se cumple en su totalidad en el nivel 3. Se cuenta con un mecanismo de chequeo de integridad sobre el contenido, el cual es proporcionado por DSpace<sup>1</sup>, el software en el que está desarrollado CIC Digital, con el cual, a través de *cronjobs*<sup>2</sup> configuradas en el servidor, se realizan chequeos periódicos sobre la integridad de todo el contenido. No obstante, estos chequeos se realizan solo sobre la copia en funcionamiento del repositorio, no se ejecutan sobre ninguna de las copias de seguridad y, si bien es posible ejecutarlas bajo demanda en respuesta a acciones específicas, por ejemplo ante la ingesta de un ítem o la modificación de los metadatos de un recurso, solo se tiene programada de manera automática la ejecución ante alguno de estos eventos, se debe realizar de forma manual para el resto. Por esto último es que se marcaron en amarillo, es decir como punto parcialmente satisfecho, dos de las recomendaciones del último nivel. Por otro lado, no se realiza un análisis de virus sobre el contenido y en consecuencia no se cuenta con un mecanismo para aislar en cuarentena a un elemento sospechoso.

Sobre la categoría “Seguridad de la información”, se cumple satisfactoriamente con el primer nivel, pues DSpace provee de un módulo de autenticación y autorización en el cual se pueden definir usuarios y grupos y distintos permisos sobre el contenido para esos usuarios/grupos. Por lo tanto, es posible controlar y definir el tipo de acceso de las personas al contenido, es posible, por ejemplo, otorgar al usuario final permiso de lectura sobre todo el contenido excepto el embargado o el privado, al usuario logueado y autenticado permiso de ingesta de ítems sobre determinadas colecciones, y al usuario administrador permiso de escritura sobre todo el contenido. A nivel de servidor

---

<sup>1</sup> <https://duraspace.org/dspace/>

<sup>2</sup> Procesos en segundo plano que se ejecutan a intervalos regulares. [https://es.wikipedia.org/wiki/Cron\\_%28Unix%29](https://es.wikipedia.org/wiki/Cron_%28Unix%29)

se realiza lo mismo con los grupos y usuarios del sistema operativo utilizado. Si bien DSpace de alguna manera documenta los permisos aplicados dentro de su módulo de administración de autorización, y lo mismo ocurre con los permisos en el servidor en donde es posible visualizarlos al ingresar a los directorios correspondientes, estos no se encuentran formalizados ni centralizados en ningún documento institucional. Por otra parte, en lo que se refiere a logs y registros de cambios, el metadato *provenance* mantiene un pequeño registro de los usuarios que participaron en la ingesta de un ítem, tanto el encargado de la revisión como el que realizó el envío; sin embargo, si un usuario administrador realiza modificaciones sobre el ítem desde la pantalla de edición para administradores, estos no quedan registrados. A su vez, en el servidor se registran los logs de acceso que provee el servidor web, pero no se revisan con regularidad, sólo en respuesta a eventos específicos, y sólo se mantienen las copias de los logs que tienen unos pocos días de antigüedad, el resto se elimina.

Con respecto a la categoría "Metadatos", se utiliza un esquema propio, mezclando elementos de distintos esquemas como dc, dcterms y un esquema propio. Todos los metadatos se encuentran inventariados en la base de datos y con dos copias de seguridad, lo que permite cumplir con el primer nivel de esta categoría. DSpace además provee de algunos metadatos administrativos y técnicos, que guardan información de la fecha de ingreso del contenido y del usuario que realizó la carga (por ejemplo el metadato dc.provenance) y la revisión del ítem, además de los metadatos descriptivos. Se realizan revisiones por parte de los usuarios administradores dedicados a todas las ingestas de ítems al repositorio para comprobar que el contenido cumpla con el formato propuesto para el perfil de metadatos de CIC Digital. También se realizan otras revisiones una vez ingresado el contenido, como tareas de curación<sup>3</sup> destinadas al control de calidad de los metadatos y distintas reformulaciones al

---

<sup>3</sup> Tareas que permiten de forma sencilla y extensible gestionar operaciones de rutina sobre el contenido en un repositorio. <https://wiki.lyrasis.org/display/DSDOC6x/Curation+System>

esquema de metadatos para agregar nuevos metadatos o modificar los existentes con el objetivo de que estos se adapten lo mejor posible al contenido ingresado. A pesar de proveer metadatos como el *provenance* que registran ciertos cambios en el contenido, DSpace no provee de metadatos que se ajusten a algún estándar de preservación existente, o que registren cambios en el ítem durante todo su ciclo de vida; en CIC Digital tampoco se ha implementado ninguno.

Por último, la categoría “Formato de archivos” es la única en la que se cumplen satisfactoriamente con todos los noveles. Desde la ingesta de los ítems se definen los formatos de archivos que se permiten cargar en el repositorio. Estos formatos se actualizan para ofrecer siempre los formatos más utilizados y con el mayor soporte dependiendo del tipo de recurso ingresado. Para el control de obsolescencia de los formatos se realizan chequeos y consultas manuales sobre los datos y si algún objeto digital es ingresado en una versión obsoleta de alguno de los formatos de ingesta, se lo transforma a la versión adecuada. Por ejemplo, en el caso de los PDF, además de realizarse la extracción automática de texto, se realizan chequeos y normalizaciones para que todos se encuentren en el formato PDF/A.

## Propuesta de mejora

Siguiendo la distribución de las recomendaciones de la matriz de NDSA, se presenta una propuesta de mejora a cada uno de los puntos no satisfechos por CIC Digital, agrupados por categoría:

### Almacenamiento y localización geográfica

Como primera medida sería deseable que el backup que está en una localización geográfica distinta (en la Facultad de Ciencias Jurídicas y Sociales de la UNLP), se encuentre además en una red distinta a la del resto de las copias. Así, si llega a haber alguna falla que afecte a toda la red de la UNLP, el backup sería todavía totalmente accesible, sin necesidad de un acceso físico al servidor en donde se guardan.

El siguiente paso en la prevención de amenazas y riesgos sobre los datos sería el almacenamiento de una de las copias de seguridad en una ciudad distinta de la que se almacenan las restantes, y que esta ciudad no sea afectada por las mismas amenazas ambientales (por ejemplo, un desastre natural) que la primera. Una solución viable para este caso sería el almacenamiento de la copia de seguridad en algún servicio en la nube, por ejemplo, Glacier Deep Archive de Amazon, un servicio de almacenamiento a largo plazo que se autoproclama durable, seguro y de bajo costo y que permite cumplir con las leyes vigentes en cuanto a privacidad y seguridad de los datos. En este servicio, los datos son almacenados en cintas y se replican en tres zonas de disponibilidad distintas dentro de la infraestructura de Amazon lo que asegura una disponibilidad mayor al 99 %.

También se debería documentar todos los procesos que participan en la creación de los backups, junto a los soportes y sistemas de almacenamiento disponibles y elaborar un plan de gestión de riesgos adecuado, que contemple casos en los que los distintos sistemas de almacenamiento y backups queden inaccesibles.

### **No alteración de archivos e integridad de los datos**

DSpace provee de una tarea de curación para el análisis de virus sobre los ítems, pero se debe ejecutar de forma manual. Se propone primero realizar un análisis sobre la tarea de curación para saber si realmente cumple su propósito y luego, si esto se cumple, automatizar su ejecución para que se ejecute cada cierto período de tiempo sobre todo el contenido del repositorio, mediante alguna cronjob. También se podría ejecutar ante cada ingesta de un nuevo ítem al repositorio, durante el proceso de revisión. Algo parecido se podría implementar para la ejecución del chequeo de integridad para que, además de ejecutarse periódicamente sobre todo el contenido y sobre un ítem específico durante su ingesta, se ejecute ante otro tipo de eventos, como la edición de un ítem, y también se debería ejecutar este chequeo sobre las copias de seguridad del contenido, para lo cual, a priori, se tendría que descomprimir el contenido de la copia de seguridad y luego ejecutar el chequeo de integridad de DSpace

desde la consola del servidor en donde se encuentre la copia, lo que implica también la instalación en ese servidor de las herramientas correspondientes (Java, DSpace, etc).

### **Seguridad de la información**

Una de las opciones para mejorar el seguimiento y el registro de quién ha realizado qué acción con qué archivo es la de la incorporación de metadatos de preservación al repositorio. Si bien se cuenta con el metadato *provenance*, este no registra todas las acciones que se efectúan sobre un ítem y su sintaxis no es la adecuada. Para la correcta inclusión de los metadatos de preservación necesarios y su correcto uso con la sintaxis correspondiente, se debería mirar la propuesta del diccionario de datos PREMIS. Por el lado del servidor, serviría como una primera medida mantener los backups de los logs de acceso por un tiempo mayor al que se mantiene actualmente, al menos se podría almacenar la información de registros de los últimos 15 días.

### **Metadatos**

Con respecto a la protección y localización de los metadatos, DSpace (y por lo tanto CIC Digital) ya provee un inventario y una clara localización del almacenamiento al guardar los metadatos en una tabla dedicada dentro de la base de datos. Esto, a su vez, también permite la separación del almacenamiento y las copias de seguridad de los metadatos en distintas localizaciones si se realizan periódicos backups de la base de datos y estos backups se trasladan a otro servidor, el cual puede estar incluso en una localización geográfica distinta al servidor de base de datos.

Si bien DSpace crea automáticamente metadatos, por ejemplo el metadato *provenance*, con información administrativa sobre quién crea el contenido, la fecha de creación y de última modificación, y quién realizó una edición sobre los metadatos de un ítem, esa información no se registra por metadato sino que se registra sobre el ítem como una unidad y tampoco se mantiene un histórico o historial de estos cambios, por lo que a veces se puede saber quién fue la última persona o la última vez que el ítem sufrió modificaciones, pero no

se puede saber sobre qué metadato se hicieron ni tampoco cuáles fueron las modificaciones previas a esa y quién las hizo. A su vez, la sintaxis utilizada para guardar la poca información de preservación con la que se cuenta, no se adecua a ningún estándar de metadatos de preservación conocido. La implementación y adopción de los correctos metadatos de preservación, en particular adaptar el perfil actual de metadatos, y modificar alguno de los esquemas utilizados, para emular la propuesta del diccionario PREMIS de Metadatos de Preservación. Si bien PREMIS es, en realidad, una propuesta de unidades semánticas (objetos, agentes, derechos y eventos) que propone un esquema jerárquico de metadatos, tal vez pueda plantearse adaptar el diccionario de datos de PREMIS a un esquema plano, utilizando alguno de los esquemas usados por el perfil de metadatos de CIC Digital, por ejemplo dcterms. Esto solucionaría los problemas de falta de información en cuanto a modificaciones sobre los metadatos y el formato con que esa información se conserva.

### **Formatos de archivos**

Se implementó una tarea de curación que chequea si los PDF de los ítems se encuentran en formato PDF/A; la ejecución periódica de esta tarea ayudaría a tener un control más automatizado sobre los formatos de los archivos. Se podría implementar de manera similar otras tareas que chequeen que el resto de los formatos se encuentren vigentes y que se pueda ejecutar esa tarea de manera automática.

### **Conclusión**

La matriz de autoevaluación propuesta por NDSA permite a un repositorio determinar su estado en materia de preservación digital, de manera sencilla y relativamente rápida. Basta con un diagnóstico y una mirada crítica sobre la documentación, implementación y políticas del repositorio por parte del personal experto para poder detectar sus fortalezas y falencias. Este esquema es ideal para instituciones o repositorios que no poseen los recursos o la

capacidad (tanto económica como de personal), de contratar una auditoría externa experta. Funciona como un primer paso de evaluación en el camino hacia la certificación en preservación, pero no es un reemplazo de los métodos existentes, como el cumplimiento de las distintas normas ISO u otros sistemas de auditoría, que son la siguiente etapa a la que debería apuntar un repositorio.

Por estas razones fue que se decidió realizar esta evaluación sobre CIC Digital, un repositorio de tamaño mediano al que nunca se le había realizado ningún análisis en términos de preservación digital. No se debe pasar por alto que NDSA Levels por su sencillez y su carácter subjetivo (por ser una autoevaluación), no provee un listado exhaustivo de las tareas de preservación y tiende a tener una visión optimista del estado del repositorio (TÉRMENS & LEIJA, 2017).

Si bien CIC Digital cumple de manera satisfactoria casi la totalidad del nivel 1, tiene varias falencias en el resto de los niveles, especialmente en lo que se refiere a la documentación de los distintos procesos y recursos de almacenamiento, planeamiento de la seguridad y riesgo, además de que el seguimiento de las modificaciones a los archivos y metadatos se podría mejorar considerablemente. Se cuenta con un buen nivel en la cantidad y calidad de las copias de seguridad, pero existe un faltante en la documentación de la localización y la forma de restaurar esas copias y la carencia de un plan integral que indique qué hacer ante cada situación de riesgo.

Esto último debería ser algo en lo que el repositorio se debería centrar y que no debería demandar un costo muy alto, tanto en tiempo como en dificultad. Quizás incluso el plan integral de riesgos es viable en el corto plazo para el repositorio, a diferencia del resto de medidas que se deberían tomar sobre las copias de seguridad, por ejemplo, la creación de otra copia de seguridad en una localización con amenaza de desastre diferente, que implicaría un costo de inversión en infraestructura o la contratación de servicios en la nube. Otras acciones a implementar que implican solo ejecuciones de tareas ya implementadas, como el análisis de virus, o el chequeo de calidad de los metadatos y el chequeo sobre el formato de los archivos, no demandarían un

esfuerzo tecnológico significativo (bastaría con la configuración de algunas cronjobs y tareas de curación) y serían realizables en el corto plazo.

## Bibliografía

- BIBLIOTECA NACIONAL DE ESPAÑA. (n.d.). Diccionario de Datos PREMIS de Metadatos de Preservación. PUBLICACIONES DE LA BIBLIOTECA NACIONAL DE ESPAÑA. <http://www.bne.es/es/Micrositios/Publicaciones/PREMIS/>
- FERRERAS-FERNÁNDEZ, T. (2010). Preservación digital en repositorios institucionales: GREDOS. Universidad de Salamanca, tesis de maestría. [https://www.researchgate.net/publication/223905922\\_Preservacion\\_digital\\_en\\_repositorios\\_institucionales\\_GREDOS](https://www.researchgate.net/publication/223905922_Preservacion_digital_en_repositorios_institucionales_GREDOS)
- LEIJA, D.; TÉRMENS, M. (2019). Traducción de Niveles de Preservación Digital NDSA 2019: Traducción al Español de Versión 2.0. APREDIG - Asociación Iberoamericana de Preservación Digital.
- LEVELS OF PRESERVATION REVISION WORKING GROUP, KUSSMANN, C., NATIONAL DIGITAL STEWARDSHIP ALLIANCE (NDSA), Graham, W., Atkins, W., Reich, A., & Walker, P. (2019, October). 2019 LOP Implementation Guide and Working Definitions. <https://osf.io/nt8u9/>
- NATIONAL DIGITAL INFORMATION INFRASTRUCTURE AND PRESERVATION PROGRAM (NDIIPP) OF THE LIBRARY OF CONGRESS. (n.d.). About the NDSA. National Digital Stewardship Alliance - Digital Library Federation. <https://ndsa.org/about/>
- NATIONAL DIGITAL INFORMATION INFRASTRUCTURE AND PRESERVATION PROGRAM (NDIIPP) OF THE LIBRARY OF CONGRESS. (n.d.-b). Levels of Digital Preservation. National Digital Stewardship Alliance - Digital Library Federation. <https://ndsa.org/publications/levels-of-digital-preservation/>
- NDSA LEVELS OF PRESERVATION ASSESSMENT SUBGROUP. (2019). Using the Levels of Digital Preservation as an Assesment Tool. <https://doi.org/10.17605/OSF.IO/OGZ98>
- TÉRMENS, M.; LEIJA, D. (2017). Auditoría de preservación digital con NDSA Levels. *Profesional de la Información*, 26(3), 447–456. <https://doi.org/10.3145/epi.2017.may.11>

## Anexo 1

Traducción al español de la matriz de preservación digital de NDSA

Área Funcional	Nivel			
	Nivel 1 - (Conocer su contenido)	Nivel 2 - (Proteger su contenido)	Nivel 3 - (Controlar su contenido)	Nivel 4 - (Mantener su contenido)
<b>Almacenamiento</b>	Tener dos copias completas en ubicaciones separadas  Documentar todos los medios de almacenamiento donde este almacenado el contenido  Poner el contenido en soportes de almacenamiento estables	Tener tres copias completas con al menos una copia en una ubicación geográfica distinta  Documentar el almacenamiento y medios de almacenamiento, indicando los recursos y las dependencias que estos requieren para funcionar	Tener al menos una copia en una ubicación geográfica con amenaza de desastre diferente a las otras copias  Tener al menos una copia en un medio de almacenamiento de diferente tipo  Rastrear la obsolescencia del almacenamiento y los medios	Tener al menos tres copias en ubicaciones geográficas distintas, cada una con una amenaza de desastre diferente  Maximizar la diversificación del almacenamiento para evitar puntos únicos de falla  Tener un plan y realizar acciones para abordar la obsolescencia del hardware, software y medios de almacenamiento
<b>Integridad</b>	Verificar que la información de integridad se ha proporcionado con el contenido  Generar información de integridad si esta no ha sido proporcionada con el contenido  Se verifica virus en todo el contenido; se aísla el contenido en cuarentena según sea necesario	Verificar la información de integridad al mover o copiar contenido  Usar bloqueadores de escritura cuando se trabaja con medios originales  Hacer una copia de seguridad de la información de integridad y almacenar una copia en una ubicación separada del contenido	Verificar la información de integridad del contenido en intervalos fijos  Documentar los procesos y resultados de verificación de información de integridad  Realizar una auditoría de la información de integridad bajo demanda	Verificar la información de integridad en respuesta a eventos o actividades específicas  Reemplazar o reparar el contenido dañado según sea necesario
<b>Control</b>	Se determinan los agentes humanos y de software que deben estar autorizados para leer, escribir, mover y eliminar contenido	Documentar a los agentes humanos y de software autorizados para leer, escribir, mover y eliminar contenido y aplicar estos	Mantener los registros (logs) y se identifican a los agentes humanos y de software que realizaron acciones sobre el contenido.	Se realizan revisiones periódicas de acciones / registros (logs) de acceso
<b>Metadatos</b>	Crear un inventario de contenido, documentando también la ubicación de almacenamiento actual de estos  Hacer una copia de respaldo del inventario y se almacena al menos una copia por separado	Almacenar suficientes metadatos para saber cuál es el contenido (esto podría incluir alguna combinación de aspectos administrativos, técnicos, descriptivos, de preservación y estructurales)	Determinar qué estándares de metadatos aplicar  Encuentra y completa los vacíos en sus metadatos para cumplir con esos estándares	Registrar las acciones de preservación asociadas con el contenido y cuándo ocurren esas acciones Implementa los estándares de metadatos elegidos
<b>Contenido</b>	Documentar los formatos de archivo y otras características de contenido esenciales, incluido cómo y cuándo fueron identificados	Verificar los formatos de archivo y otras características de contenido esenciales  Establecer relaciones con los creadores de contenido para fomentar la elección sostenible de archivos	Monitorear la obsolescencia y los cambios en las tecnologías de las que depende el contenido	Realizar migraciones, normalizaciones, emulación y actividades similares que garanticen el acceso al contenido

Traducción de Niveles de Preservación Digital NDSA 2019: traducción al español de Versión 2.0. (LEIJA & TÉRMENS, 2019).

# Experiencias en la migración de un repositorio institucional a DSpace 7

## **PABLO CÉSAR DE ALBUQUERQUE**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[pablo@sedici.unlp.edu.ar](mailto:pablo@sedici.unlp.edu.ar)

## **SANTIAGO TETTAMANTI**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[santit@sedici.unlp.edu.ar](mailto:santit@sedici.unlp.edu.ar)

## **GONZALO LUJÁN VILLARREAL**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[gonzalo@prebi.unlp.edu.ar](mailto:gonzalo@prebi.unlp.edu.ar)

## **ARIEL JORGE LIRA**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[alira@sedici.unlp.edu.ar](mailto:alira@sedici.unlp.edu.ar)

## **MARISA RAQUEL DE GIUSTI**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[marisa.degiusti@sedici.unlp.edu.ar](mailto:marisa.degiusti@sedici.unlp.edu.ar)

## **RESUMEN**

En este trabajo se cuenta la experiencia obtenida a partir de la migración del repositorio institucional de la Comisión de Investigaciones Científicas (CIC), denominado CIC-DIGITAL, basado en

el software de repositorios DSpace a su última versión (DSpace 7). Se realiza un análisis de este software, mencionando las diferencias arquitecturales y funcionales con respecto a sus versiones anteriores. Luego se explica la estrategia seguida a lo largo del proceso de migración junto con los desafíos encontrados durante el mismo. Finalmente, se reflexiona sobre las distintas estrategias de adopción de este software para los distintos escenarios en los que puede encontrarse un repositorio institucional que pretende encarar una migración de características similares.

#### **PALABRAS CLAVE**

Repositorios digitales; DSpace 7; migración de software.

*Digital repositories, DSpace 7; software migration.*

## **Introducción**

La versión 7 de DSpace<sup>1</sup> difiere de sus antecesoras en varios aspectos, siendo uno de los más importantes el cambio en su arquitectura y la tecnología usada en la interfaz del usuario. A partir de la versión 7, el software se divide en dos proyectos independientes, un backend y un frontend que se comunican a través de una nueva API REST que expone toda la funcionalidad del repositorio. Si bien esta versión trae nuevas ventajas y facilita el desarrollo de nuevos servicios gracias al desacoplamiento del backend y la funcionalidad expuesta por la API REST, al día de hoy no toda la funcionalidad de las versiones anteriores está implementada en DSpace 7. Es por esto último que la adopción de esta nueva versión puede ser más o menos dificultosa, dependiendo del grado de desarrollo de los repositorios digitales en cada institución.

---

<sup>1</sup> DSpace 7: <https://duraspace.org/dspace/dspace-7/>

En este trabajo se busca compartir la experiencia obtenida por el equipo de desarrollo del repositorio CIC-DIGITAL, responsable del desarrollo del repositorio SEDICI de la Universidad Nacional de La Plata (UNLP), a partir de la migración del repositorio a DSpace y exponer las dificultades encontradas para que otras instituciones interesadas en utilizar esta versión de DSpace, cuenten con más herramientas a la hora de definir una estrategia de adopción de este software, ya sea a partir de una migración o de la implementación desde cero.

## Arquitectura de DSpace

### DSpace 6 y versiones anteriores

DSpace es el software para repositorios más utilizado en la actualidad: es utilizado por aproximadamente el 39 % de los repositorios (OpenDOAR Statistics - v2.sherpa, 2022). La versión 6 y las anteriores estaban compuestas de varios proyectos, gestionados a través de la herramienta Maven<sup>2</sup>, entre los cuales dos de ellos eran utilizados para implementar la interfaz del usuario. Estos proyectos son dspace-jspui y dspace-xmlui, desarrollados con Java Servlets y el framework Apache Cocoon, respectivamente. En la IMAGEN 1 se puede apreciar, dentro de la línea punteada roja, las seis aplicaciones correspondientes a la capa de aplicación de DSpace, entre las cuales figuran JSPUI y XMLUI. Es importante observar que el desarrollo del frontend y del backend se encuentra dentro de un mismo proyecto.

---

<sup>2</sup> Maven: <https://maven.apache.org/>

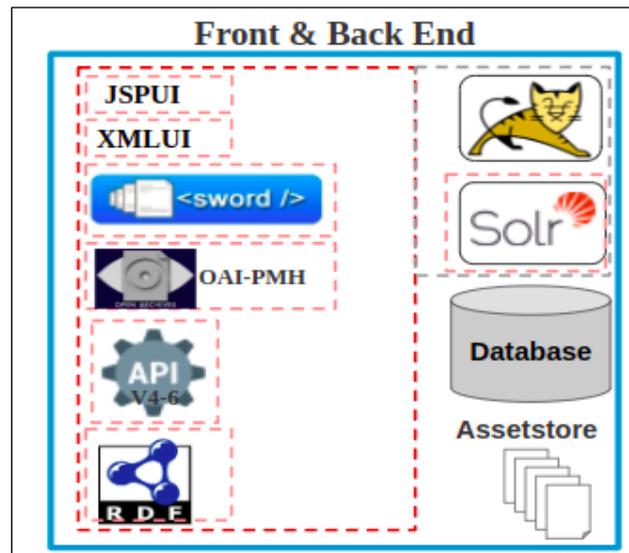


IMAGEN 1. Arquitectura de DSpace 6 (DONOHUE *et al.*, 2021)

En este punto, cabe remarcar que todas las tecnologías usadas por JSPUI y por XMLUI están, o bien en desuso (servlet, JSP, Java 8) o bien abandonadas (Apache Cocoon). Tal es así, que el propio equipo de desarrollo de DSpace decidió discontinuar el soporte de ambas herramientas.

### DSpace 7

DSpace lanzó la versión 7.0 en agosto del 2021. Esta versión promete mantener la misma funcionalidad que las versiones anteriores y además agregar nuevas características como, por ejemplo, el manejo de entidades configurables y las relaciones entre ellas. Además de la ampliación de su funcionalidad, la publicación de esta nueva versión introdujo también un cambio de tecnologías y de la estructura general del proyecto. Por el lado de la interfaz de usuario o frontend, se desarrolló una aplicación web en Angular<sup>3</sup> y de manera desacoplada del resto del sistema; esta decisión implicó la creación de una nueva interfaz de usuario, así como también la eliminación de los módulos de frontend de las anteriores versiones. Por el lado del sistema de gestión o backend, se dejó de lado el uso del software Cocoon (ya deprecado)

<sup>3</sup> Angular: <https://angular.io/>

para pasar a utilizar Spring de forma más integral, y se construyó una nueva API REST que brinda acceso a toda la funcionalidad del repositorio<sup>4</sup>.

El salto hacia nuevas tecnologías propuesto en esta versión de DSpace, sumado a la incorporación de nueva funcionalidad, a una API REST que brinda acceso a toda funcionalidad del sistema, y a la posibilidad de implementar una interfaz de usuario más amigable, hace que sea muy conveniente para un repositorio desarrollado en DSpace 6 o en versiones anteriores migrar hacia DSpace 7. Sin embargo, esta migración no siempre puede realizarse de manera directa: si bien gran parte de la lógica de negocios se mantuvo entre versiones, se deben readaptar las personalizaciones realizadas a la nueva estructura de API REST y frontend desacoplado, y se deben trasladar las personalizaciones en la interfaz de usuario hacia una nueva tecnología o bien realizar su rediseño.

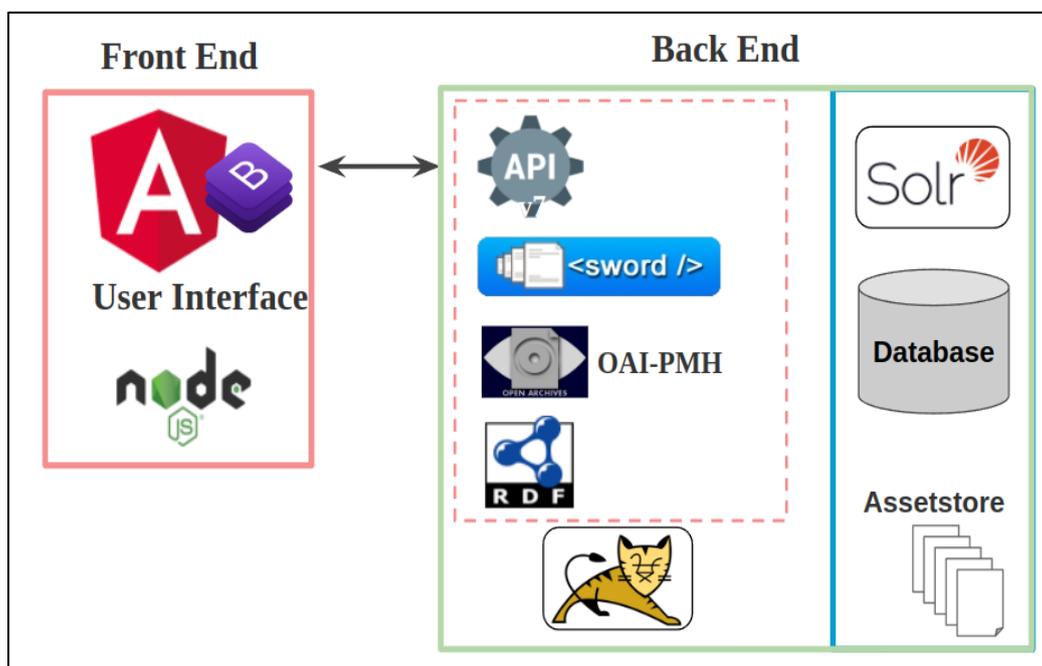


IMAGEN 2. Comunicación entre componentes de la arquitectura de DSpace 7 (basado en DONOHUE et al., 2021)

<sup>4</sup> Si bien las versiones anteriores de DSpace ya disponían de una API REST, esta brindaba acceso a un conjunto muy acotado de funcionalidades del repositorio, y su desarrollo estaba enfocado en la integración de los repositorios con sistemas de terceros o herramientas externas.

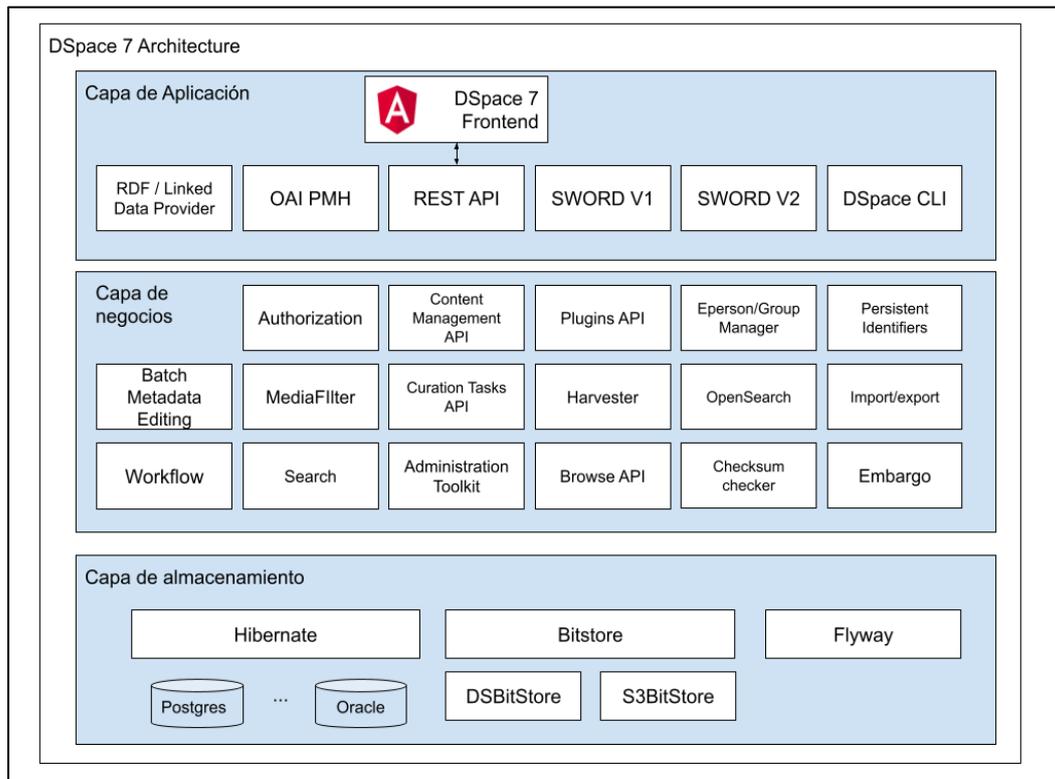


IMAGEN 3. Organización de componentes de la arquitectura de DSpace 7

### Beneficios de la nueva versión de DSpace

- Desacoplamiento del backend en un nuevo módulo, que expone toda su funcionalidad a través de una API REST, que a su vez permite integrar más fácilmente nuevos servicios.
- Implementación de una interfaz moderna basada en componentes reutilizables, dejando de lado tecnologías que han quedado sin soporte ni actualizaciones.
- Facilidad para integrar nuevos servicios, gracias al desacoplamiento antes mencionado. Esta mejora aumenta el grado de interoperabilidad del repositorio con otros sistemas a partir del uso de una nueva API REST más completa que la existente en versiones anteriores de Dspace.
- Posibilidad de extender el modelo para orientarlo hacia un sistema CRIS (ASSERSON & JEFFERY, 2010) que gestione otros tipos de entidades, como “Personas”, “Organizaciones”, “Proyectos”, etcétera.

## Funcionalidad nueva y pendiente a la versión 7.2.1

### Configurable Entities

Una de las nuevas funcionalidades de la versión 7 de DSpace recae en la posibilidad de expandir el modelo subyacente para poder representar no solo recursos, como artículos, tesis o presentaciones en congresos, sino que ahora es posible representar entidades que forman parte de un modelo CRIS, como lo son las personas, organizaciones, proyectos, revistas, etc. A estas nuevas entidades se las llama *Configurable Entities* (DSpace 7 - Configurable Entities, n.d.) y no sólo modelan entidades, sino también las relaciones que pueden existir entre ellas. En la IMAGEN 4 se pueden observar cuatro entidades (Publicación, Proyecto, Persona y Organización) así como las relaciones que se definieron entre ellas.

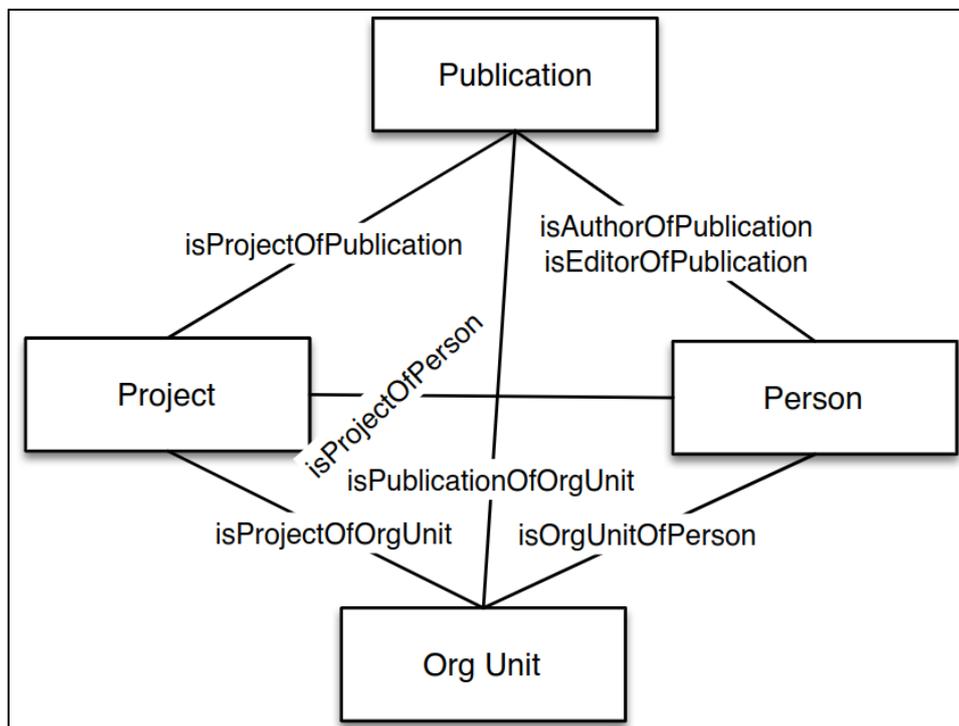


IMAGEN 4. Tomada de BOSMAN et al., 2021

La adición de las Configurable Entities permite gestionar nuevos tipos de recursos y crear páginas para cada una de las instancias de estas entidades,

así como se hace con los Ítems. Un ejemplo del potencial de esta nueva funcionalidad es que es posible crear perfiles de autor a partir de los valores almacenados en los metadatos que describen quién es el creador de un determinado recurso y gestionarlos dentro de DSpace y no a través de un sistema de gestión de vocabularios controlados externo como se hace en CIC-DIGITAL (DE ALBUQUERQUE, 2018).

### **DSpace Angular**

Para la interfaz de usuario, DSpace 7 ofrece un frontend basado en Angular que reemplaza a las interfaces XMLUI y JSPUI. Angular es un framework de código abierto orientado al desarrollo de aplicaciones frontend basado en TypeScript. Este framework permite definir componentes reutilizables, a los que se puede dotar con funcionalidades que enriquezcan la experiencia del usuario, logrando así interfaces potentes y flexibles. Además de basarse en el lenguaje TypeScript, mucho más expresivo y seguro que Javascript, Angular se destaca por su alto grado de adopción y por la activa y creciente comunidad de desarrolladores.

El proyecto DSpace Angular provee tres directorios de temas, entre los cuales uno de ellos define los componentes principales del frontend, mientras que los otros dos permiten su personalización. Estos temas son:

- Base Theme
- DSpace Theme
- Custom Theme

La definición de estos temas es de suma importancia ya que lo que se busca es una forma de personalizar el frontend sin tocar los componentes principales y evitar conflictos con el desarrollo principal del proyecto.

## Migración

### Características de CIC-DIGITAL

El repositorio institucional de la Comisión de Investigaciones Científicas de la provincia de Buenos Aires (CIC), denominado CIC-DIGITAL, fue creado a finales de 2014. Además de cumplir el rol de archivo digital institucional, el repositorio asiste a los investigadores, personal de apoyo y becarios de la CIC para que puedan disponer en Internet su producción, maximizando su difusión e impacto, a la vez que protege a largo plazo sus obras. CIC-DIGITAL funciona bajo la dirección <https://digital.cic.gba.gob.ar> y a la fecha cuenta con casi 11.000 publicaciones en acceso abierto. Desde 2015 está adherido al SNRD<sup>5</sup> y es cosechado en el portal nacional de repositorios digitales.

CIC-DIGITAL fue originalmente desarrollado sobre DSpace 4 y actualizado en repetidas ocasiones hasta llegar a la versión 6.4-snapshot, es decir, a la última versión disponible en la rama 6 y previa a DSpace 7. Si bien 6.4-snapshot no es un *release* formal y estable, esta versión contiene modificaciones, parches de seguridad y arreglo de errores de importancia para el correcto funcionamiento del repositorio. CIC-DIGITAL posee además una serie de personalizaciones sobre Dspace, realizadas a partir de distintos requerimientos que surgieron en algunos casos por necesidades detectadas por los administradores durante el uso del repositorio a lo largo de los años, y en otros casos a partir de las políticas propias de la institución. El stack de tecnologías utilizado en CIC-DIGITAL es el estándar de DSpace, es decir, se utiliza Solr como indexador de texto para búsquedas y estadísticas y PostgreSQL como base de datos. La interfaz de usuario está desarrollada íntegramente con XMLUI, y para la personalización de dicha interfaz se creó un tema propio, denominado cicba<sup>6</sup>, el cual es una extensión al tema Mirage que es el tema por defecto de DSpace 6.

---

<sup>5</sup> Sistema Nacional de Repositorios Digitales, Ministerio de Ciencia, Tecnología e Innovación: <https://repositoriosdigitales.mincyt.gob.ar/>

<sup>6</sup> Github CIC-DIGITAL: <https://github.com/CICBA/DSpace/releases/tag/dspace-cic-6.4.3>

Además de las adaptaciones realizadas sobre el código de DSpace, CIC-DIGITAL cuenta también con sistema de autoridades implementado en el sistema de gestión de contenidos Drupal<sup>7</sup> en versión 9. La interconexión entre Drupal y DSpace se realiza mediante una API REST y una serie de conectores a medida desarrollados sobre CIC-DIGITAL para tal fin.

### **Definición de un plan de migración**

El proceso de migración se puede dividir en dos: la migración del frontend y la migración del backend. Por el lado del backend, la migración<sup>8</sup> consistió en una serie de pasos que implican la actualización del código de CIC-DIGITAL hacia la nueva versión. Esto implicó incorporar los cambios realizados por DSpace en su proyecto principal de desarrollo de la versión 7 con el proyecto de CIC-DIGITAL. El frontend, por su parte, requirió de una iteración continua de diseño de la nueva interfaz<sup>9</sup>, implementación de lo propuesto y pruebas (testing) de lo realizado, hasta llegar al punto de tener una interfaz de usuario usable y funcional, apta para ser utilizada por los usuarios del repositorio.

### **Migración de backend**

La migración del backend consistió en lo siguiente:

1. Como primer paso, antes de realizar la incorporación (*merge*) de la versión a actualizar de DSpace, se realizaron análisis y tareas de preparación para que el proceso de merge sea lo más sencillo posible:
  - Se realizó un relevamiento de cambios en la nueva versión de DSpace. Esto sirvió, además de para conocer la funcionalidad disponible en la nueva versión, para saber qué cambios iba a sufrir el código proyecto de CIC Digital.

---

<sup>7</sup> Drupal: <https://www.drupal.org/>

<sup>8</sup> Migración a DSpace7: <https://github.com/CICBA/DSpace/pull/55>

<sup>9</sup> CICBA/DSpace-Angular: <https://github.com/CICBA/DSpace-angular>

- Se analizaron los *fixes* (correcciones) superpuestos con los resueltos en CIC-Digital. Al detectar correcciones de errores realizados en paralelo en el desarrollo de DSpace 7 y en CIC-Digital (sean correcciones propias o incorporadas desde la versión 6), se pudieron prever potenciales puntos de conflicto al realizar el *merge*.
- Se borraron proyectos que no ya no forman parte del desarrollo de DSpace 7. Con el cambio de arquitectura en la aplicación que propone DSpace 7 se descontinuaron varios de los proyectos que se mantienen en uso hasta la versión 6.
  - I. dspace-jspui
  - II. dspace-xmlui
  - III. dspace-solr, ya que en la nueva versión el indexador se instala de manera desacoplada al resto de la aplicación
- 2. Se incorporó (*merge*) el código de DSpace correspondiente a la última versión estable de la rama 7.x (en el caso de CIC-DIGITAL, la versión estable al momento de hacer la migración era la dspace-7.1<sup>10</sup>). En este punto se identificaron algunos conflictos entre el nuevo código de DSpace y el código de CIC-DIGITAL debido a algunas personalizaciones desarrolladas.
- 3. Luego se buscó diferenciar los conflictos a partir del reporte de git. Con esta herramienta fue posible saber si el conflicto fue originado por código proveniente de CIC-DIGITAL o código proveniente del desarrollo de DSpace. Luego, se repasaron los conflictos resultantes de la integración, se analizaron los cambios implementados en CIC-DIGITAL y se definieron qué modificaciones deben conservarse.
- 4. Una vez analizados los conflictos, se procedió a resolverlos a partir de la diferenciación propuesta por git. La resolución de cada conflicto fue específica para cada caso en particular, aunque algunos se pudieron solucionar de forma genérica. Por ejemplo, un caso de resolución

---

<sup>10</sup> DSpace 7.1: <https://github.com/DSpace/DSpace/releases/tag/dspace-7.1>

genérica de conflictos se dió en archivos modificados tanto por DSpace como por CIC-DIGITAL, pero que no tenían en su contenido personalizaciones propias de CIC-DIGITAL, sino que contenían sólo modificaciones propias de DSpace. En esos archivos el conflicto surgía porque se habían realizado cambios en la versión 6 que no se incorporaron en la versión 7 de DSpace, en esos casos se determinó quedarse con los cambios realizados en la versión 7.

5. Una vez que se resolvieron los conflictos, hubo que adaptar las personalizaciones realizadas en CIC-DIGITAL para que sean consistentes con los cambios de esta nueva versión y el código pudiese compilar en DSpace 7. Por ejemplo, en el módulo de autoridades, el mecanismo de selección de una autoridad fue modificado, esto hizo necesario adaptar los conectores<sup>11</sup> definidos en CIC-DIGITAL a estas modificaciones y así para consumir del sistema de gestión de autoridades externo que se utiliza para gestionar investigadores de la CIC, instituciones y taxonomías usadas para describir items.

### **Migración de datos**

DSpace posee una herramienta para realizar migraciones automáticas entre las distintas versiones. Esta herramienta se utiliza desde la interfaz CLI de DSpace y hace uso del software Flyway, muy utilizado para realizar control de versiones en bases de datos. Para ejecutarla se debe hacerlo sobre una instancia de DSpace en funcionamiento que haya sido actualizada de una versión de DSpace a otra más nueva; su ejecución modifica el modelo de la base de datos relacional para adaptarlo a los cambios introducidos en la nueva versión. Para la migración de la base de datos de CIC-DIGITAL, una vez que se migró todo el código se realizó una copia de la base de datos del repositorio, para no utilizar la base de datos original, a modo de prueba, y luego se utilizó la herramienta de migración para llevar la base de datos a la última versión

---

<sup>11</sup> <https://github.com/CICBA/DSpace/commit/8acf8c8c52e9748f816b0d2ed0f5ae26da9ca0a9>

disponible en DSpace 7. La ejecución de la herramienta resultó exitosa, sin mayores inconvenientes. Una vez realizadas las pruebas pertinentes en el sistema migrado, se procedió a realizar la migración de la base de datos en producción.

### **Funcionalidad faltante en *backend***

Una de las funcionalidades que no se implementaron en DSpace 7.0 fue la de Request a Copy, que permite a un usuario contactar directamente al autor de un ítem en estado embargado con el fin de solicitar su acceso. Esta funcionalidad forma parte de la versión 6 de CIC-DIGITAL, pero al no ser un comportamiento bloqueante para su uso no se volvió a implementar. Afortunadamente, este módulo sí se encuentra presente en la versión 7.1<sup>12</sup>, por lo que su configuración queda planteada para más adelante.

### **Migración *frontend***

#### **Implementación de tema propio en Frontend Angular**

Para el *frontend* se debió rehacer la interfaz de usuario, lo que requirió de numerosas acciones de diseño, implementación y pruebas hasta obtener la interfaz de usuario finalizada. A modo descriptivo se enumeran algunas de estas acciones:

- Armado de *mockups* (maquetas) y prototipos de todo el sitio: priorizando página del inicio, vista de un ítem, página de búsqueda y la visualización de los centros CIC. Se eligieron estas páginas en particular por ser las más visitadas por los usuarios y las que definen la identidad estética del sitio.
- Implementación de lo diseñado sobre DSpace.
- Pruebas: el desarrollo en progreso se puso en funcionamiento en un servidor de pruebas para que los usuarios administradores del repositorio testeen y reporten los cambios y nuevas funcionalidades.

---

<sup>12</sup> <https://wiki.lyrasis.org/display/DSDOC7x/Request+a+Copy>

El proceso fue iterativo, y en cada ciclo se fue refinando cada vez más el detalle del diseño. Las primeras iteraciones se enfocaron más en la estructura general, colores y tipografía de las páginas, las últimas se enfocaron en el detalle de diseño y funcionalidad de cada sección y elemento de cada página en particular.

Además de realizar el rediseño de la interfaz de usuario, también se realizó un análisis funcional del *frontend*, detectando tanto funcionalidad faltante que existe en CIC-DIGITAL pero que no está implementada por parte de DSpace, como errores en el software que no permitía un correcto funcionamiento. Para el caso de la funcionalidad propia de CIC-DIGITAL, algunos cambios se decidieron implementar en la nueva versión, a otros se los pospuso para versiones futuras, y a otros se los descartó directamente. Estas decisiones dependieron de la importancia y el uso que tienen en el repositorio las funcionalidades implementadas por esas personalizaciones. Para el caso de los errores en el software, se procedió a arreglarlos y reportarlos a DSpace junto con la propuesta de resolución del mismo.

### **Funcionalidad faltante en *frontend***

Una de las principales funcionalidades ausentes en la versión 7.0 fue la del *type-bind*. Este funcionalidad asocia un conjunto de metadatos a un tipo de recurso en particular, de esta forma se pueden crear formularios más específicos que contengan sólo la información relevante para un tipo de recurso en particular, como puede ser el metadato que representa el director de una tesis (que sólo debería aplicar para el tipo de recurso de tesis). Para poder seguir con la migración se desarrolló una solución temporal, sin embargo al día de la fecha ya existe una solución<sup>13</sup>, que aún no ha sido incorporada al proyecto, por parte de otros usuarios que crearon una solución más general del problema.

---

<sup>13</sup> <https://github.com/DSpace/dspace-angular/pull/1529>

## Desafíos encontrados en la migración de D6 a D7

### Frontend

Angular es un *framework* que permite definir componentes reutilizables en toda la aplicación; sin embargo, no tiene definido nativamente un procedimiento para definir temas que hereden funcionalidad de otros temas. Por ende, el primer desafío encontrado fue identificar en DSpace la manera en que debía definirse un tema principal para toda la aplicación a partir del cual se puedan personalizar los componentes que conforman el frontend. A lo largo de las versiones beta de DSpace 7, se ofrecía un método de adaptación del *frontend* a partir de un tema llamado mantis, pero dicho método no terminó de ofrecer una solución suficientemente genérica. En la versión 7.0 de DSpace, se ofrece una solución que reemplaza mantis<sup>14</sup> en la que se define un conjunto de componentes sin funcionalidad, que sólo hacen referencia a los componentes principales de DSpace, y que son fácilmente extensibles. Sin embargo esta solución tampoco resultó lo suficientemente genérica como para personalizar todos los componentes sin modificar aspectos centrales de la aplicación. Este tipo de modificaciones deberían ser evitables, ya que al modificar estos componentes del núcleo de la aplicación pueden llevar a posibles conflictos con el desarrollo principal de DSpace.

## Reflexiones sobre el proceso de migración

### Posibles estrategias para nuevos repositorios

Como se ha visto anteriormente, la migración de DSpace 6 a 7 puede requerir de una gran inversión de tiempo y recursos técnicos, que se acentúa más cuando se trata de implementaciones con un gran número de personalizaciones. Si bien es verdad que no toda la funcionalidad de las versiones anteriores están disponibles al día de hoy (en la versión 7.2.1), y es

---

<sup>14</sup> <https://github.com/DSpace/dspace-angular/pull/1063>

tentador para una institución que no tiene un repositorio aún funcionando hacer uso de esas versiones, a medida que se vaya personalizando la instalación, el salto a la versión 7 va a ser cada vez más costoso. Esto, sumado a que es poco probable que un repositorio nuevo utilice todas las funcionalidades de DSpace (ya que la adaptación al uso del software y su posterior personalización es un proceso que se realiza a lo largo del tiempo), podría indicar que la adopción de la versión 7 sea tal vez la aproximación indicada. Con esta adopción temprana una institución no sólo se asegura tener un stack de tecnologías actualizado, sino que también tiene la posibilidad de integrarse al desarrollo del código de DSpace, aportando en base a la experiencia que se va generando a lo largo del desarrollo.

### **Posibles estrategias para repositorios existentes**

En caso de que se tenga un DSpace 6 y una institución quiera dar el salto a DSpace 7 la estrategia que se puede utilizar para migrar puede ser similar a la usada en CIC-DIGITAL. En primer lugar habría que identificar las personalizaciones realizadas y de ser posibles priorizarlas. Una vez que se incorpore el código de la versión 7 a una rama de desarrollo, los conflictos resultantes van a indicar la cantidad de trabajo que puede conllevar la migración.

### **Posibles estrategias para repositorios muy personalizados**

En caso de que los conflictos sean numerosos y afecten el código de las personalizaciones incorporadas a lo largo del desarrollo, una buena opción es repensar el desarrollo para arrancar desde una instalación limpia. Si bien esta opción parece demandar un gran esfuerzo, si se tiene en cuenta que como mínimo será necesario desarrollar todo el *frontend* desde cero, quizás pensar en un desarrollo integral que incluya tanto *frontend* como *backend* puede ser una buena opción.

## Conclusiones

En relación a la nueva arquitectura propuesta, si bien en las versiones 6 y anteriores existe una API REST que permite una interacción con el repositorio, no toda su funcionalidad podía ser usada desde esta API. Esta carencia hacía que la implementación de un *frontend*, que pudiese reemplazar a las interfaces de usuario XMLUI y JSPUI, fuese inviable. Sin embargo, los esfuerzos que se han volcado en el desarrollo de la nueva API REST de DSpace 7 no sólo hacen que finalmente sea posible implementar un *frontend* desacoplado de la capa de negocios, como es el caso del proyecto DSpace Angular, sino que permite incorporar nuevos servicios o desarrollar nuevas aplicaciones que interactúen con los repositorios. Esta característica es un punto muy importante como base para el desarrollo de Next Generation Repositories (NGR) (CONFEDERATION OF OPEN ACCESS REPOSITORIES *et al.*, 2017), una iniciativa que busca posicionar a los repositorios digitales como un proveedor de servicios más que como un depósito de producción científica.

En relación a las entidades configurables, cabe resaltar que la posibilidad de extender el modelo subyacente de DSpace es sin dudas uno de los puntos más interesantes de esta última versión. Con esta funcionalidad es posible gestionar entidades propias de un modelo CRIS como personas, proyectos, organizaciones, etc., que tienen metadatos acordes a cada entidad. Por ejemplo, los metadatos asociados a las entidades *person* se basan por defecto en el esquema Person de Schema.org ([schema-person-types.xml](#)). De la misma forma que ocurre cuando un Ítem es creado en DSpace, al momento de crear una entidad *person*, se creará también un página que agrupe sus metadatos de manera tal que sea posible crear perfiles de autores para aquellos que sean considerados más relevantes para la institución, ofreciéndoles el servicio (MANZUR & TETTAMANTI, 2021).

El desacoplamiento de la API REST permite implementar nuevos sistemas que no solo consuman los datos del repositorio, sino que también permite realizar tareas de gestión de usuarios, permisos y otras funcionalidades. Con

esto en mente es posible crear nuevas aplicaciones como por ejemplo *frontends* dedicados a usuarios o tareas más específicas, como puede ser un sistema que gestione importaciones de ítems en masa, provenientes de distintas fuentes, que implique realizar transformaciones en los datos que deban ser supervisadas por un humano.

En resumen, DSpace 7 ofrece un avance tecnológico en comparación con sus versiones anteriores, y los beneficios que aporta su adopción van a incrementarse a medida que su desarrollo continúe y se termine de implementar la funcionalidad faltante y de depurar los errores propios de un software que realizó un cambio arquitectural importante e incorporó nuevas tecnologías. Sin embargo, esta adopción puede resultar compleja y requerir una cierta dedicación de tiempo tanto en la implementación del software, como en la capacitación para su uso y su posterior personalización.

De acuerdo a la experiencia obtenida durante el proceso de migración de CIC-Digital, antes de tomar la decisión de migrar un repositorio estable en DSpace 6 o en DSpace 5 a la nueva versión se debe analizar detenidamente la situación de ese repositorio en lo que respecta al grado de personalización de su instalación de DSpace y en particular al personal disponible y sus capacidades técnicas. Un repositorio con un alto grado de personalización implica un proceso de migración muy complejo debido a la necesidad de adaptar los cambios realizados a la nueva versión. En ese caso lo recomendable quizás es la implementación de un DSpace 7 “limpio”, y luego desarrollar las personalizaciones sobre esa instalación. Si en cambio el grado de personalización es menor o nulo, la incorporación de los cambios de DSpace 7 al proyecto del repositorio pareciera ser la opción correcta.

En ambos casos se recomienda encarar el proceso de migración solo si se cuenta con personal con las capacidades técnicas necesarias para afrontar un desarrollo basado en las tecnologías sobre las que está implementada esta nueva versión de DSpace. En caso de no disponer de ese personal técnico y que sea necesario capacitar al equipo de desarrollo, se debe tener en cuenta que el costo de aprendizaje de las tecnologías usadas (Docker, Angular, Java, REST,

Spring, Hibernate, etc.) puede ser bastante alto. La migración podría implicar meses de trabajo, y si no se cuenta con un equipo técnico apto para realizarla podría ser inviable. Sin embargo, las instituciones deben poner en la balanza el costo del recurso humano versus el beneficio que se puede lograr usando un software de este tipo, como ser mejores servicios para investigadores, mejora del impacto académico, adopción del modelo de ciencia abierta y más.

## Bibliografía

- ASSERSON, A., & JEFFERY, K. (2010). CRIS and Institutional Repositories. *Data Science Journal*, adpub, 1004290223-1004290223. <https://doi.org/10.2481/dsj.CRIS3>
- CONFEDERATION OF OPEN ACCESS REPOSITORIES, RODRIGUES, E., & SHEARER, K. (2017). Next Generation Repositories: Behaviours and Technical Recommendations of the COAR Next Generation Repositories Working Group. Copyright, Fair Use, Scholarly Communication, etc. <https://digitalcommons.unl.edu/scholcom/64>
- DE ALBUQUERQUE, P. C. (2018). Soporte de vocabularios controlados y autoridades en repositorios digitales [Tesis, Universidad Nacional de La Plata]. <https://digital.cic.gba.gob.ar/handle/11746/8621>
- DONOHUE, T., LOWEL, A., & BOLLINI, A. (2021, junio 7). Getting Started with DSpace 7.0: Basic Training. *Open Repositories 2021 (OR 2021)*, Virtual. <https://doi.org/10.5281/zenodo.4908060>
- DSPACE 7—Configurable Entities. (s. f.). Google Docs. [https://docs.google.com/document/d/1X0XsppZYOtPtbmq7yXwmu7FbMAfLxxOCONbw0\\_rl7jY/edit?usp=embed\\_facebook](https://docs.google.com/document/d/1X0XsppZYOtPtbmq7yXwmu7FbMAfLxxOCONbw0_rl7jY/edit?usp=embed_facebook)
- MANZUR, E., & TETTAMANTI, S. (2021). Desarrollo de servicios basados en perfiles académicos normalizados para autores de repositorios institucionales [Tesis, Facultad de Informática]. <http://sedici.unlp.edu.ar/handle/10915/122513>
- OPENDOAR Statistics—V2.sherpa. (2022). [https://v2.sherpa.ac.uk/view/repository\\_visualisations/1.html](https://v2.sherpa.ac.uk/view/repository_visualisations/1.html)

# WebID como base para el desarrollo de una marca personal en repositorios institucionales

## **PABLO CÉSAR DE ALBUQUERQUE**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[pablo@sedici.unlp.edu.ar](mailto:pablo@sedici.unlp.edu.ar)

## **GONZALO LUJÁN VILLARREAL**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI-Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) (Argentina)

[gonzalo@prebi.unlp.edu.ar](mailto:gonzalo@prebi.unlp.edu.ar)

## **MARISA RAQUEL DE GIUSTI**

PREBI-SEDICI, Universidad Nacional de La Plata (UNLP) (Argentina)

CESGI, Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

[marisa.degiusti@sedici.unlp.edu.ar](mailto:marisa.degiusti@sedici.unlp.edu.ar)

## **RESUMEN**

Los distintos actores del ámbito académico y científico producen una gran cantidad de documentos que reflejan el avance y los resultados de sus líneas de trabajo: artículos, tesis, datasets, etc. Estos recursos se difunden a través de diversas plataformas en línea, como repositorios digitales, portales de revistas y sitios de congresos, entre otros. Para que dichas plataformas puedan hacer referencia a autores definidos en otros sistemas es necesario poder identificarlos de forma unívoca. Es por eso que los autores acuden al uso de identificadores controlados por terceros (por ejemplo, ORCID y ResearcherID). El uso de identificadores de autores no es una práctica obligatoria aunque es

recomendable ya que esto favorece la interoperabilidad entre plataformas. Sin embargo, el hecho de que sea un tercero quien gestione la información de un autor atenta contra la participación autónoma, ya que este debe ajustarse a los términos y políticas de terceros para difundir su producción. En este trabajo se exploran alternativas, que se complementan con las plataformas ya establecidas, y que le brindan al autor y a las instituciones académicas la posibilidad de registrar, describir y administrar sus propias identidades sin depender de servicios de terceros.

#### **PALABRAS CLAVE**

WebID; perfiles de autores; marca personal; repositorios digitales.  
*WebID; author profiles; personal branding; digital repositories.*

## **Objetivo**

El objetivo de este trabajo es describir mecanismos y protocolos que puedan brindarle a un autor la posibilidad de crear un perfil gestionado por él mismo en un espacio propio, con la capacidad de exponer sus datos en otros sitios, y utilizar dicho espacio como base para el desarrollo de una marca personal en entornos académicos. Para esto se describirán los protocolos WebID y WebID Profile, propuestos por el Grupo de Trabajo de Repositorios de Próxima Generación (Next Generation Repository Working Group, NGRWG)<sup>1</sup>, y diseñados para identificar entidades como personas, empresas, organizaciones, entre otros, basándose en los principios propuestos por Linked Data.

---

<sup>1</sup> Web con documentación técnica sobre NGRWG: <https://ngr.coar-repositories.org/technology/webid/>

## Organización del texto

La primera parte de esta ponencia define el concepto de marca personal y resalta la importancia de otorgar a los autores el control sobre ella. Se analiza también la forma en que los autores suelen utilizar distintos servicios para poder identificarse en la web y dar a conocer su obra. La segunda parte de la ponencia introduce los conceptos de WebID y WebID Profile, y propone una forma de incorporar estas tecnologías en el ámbito de un repositorio institucional. Cómo se indicará más adelante, la incorporación de WebID en este ambiente resulta de gran utilidad a la hora de distribuir el control de las redes de repositorios.

## Marca personal en entornos académicos

El concepto de *marca personal* fue utilizado por primera vez por TOM PETERS en 1997 en su artículo “The Brand Called You” y si bien es un concepto que se origina a partir de las teorías del *branding* corporativo<sup>2</sup>, en las que si la marca supone un medio de transmisión de valores e identidad, todos los individuos tienen una marca personal, sobre la cual se construye la reputación, que es lo que los demás piensan de uno (BROWN, 2010). En palabras de GONZÁLEZ-SOLAR (2018), “todas las personas tienen una marca personal así como la capacidad de gestionarla, hacerlo de forma consciente es la fórmula para obtener sus beneficios y evitar los resultados indeseados. Tomar el control de la marca personal facilita los procesos de reputación influyendo de manera positiva en las percepciones y mejorando las posibilidades de destacar como profesional, no hacerlo permite que sean terceros quienes la configuren con sus opiniones”.

---

<sup>2</sup> “Podemos definir el *branding* corporativo como la representación mental que el público tiene de un objeto, en este caso, empresa o entidad, o la idea que otros tienen de la empresa.” Recuperado de <https://www.esic.edu/rethink/marketing-y-comunicacion/branding-corporativo-en-que-consiste-y-cual-es-su-importancia>

Hoy en día los investigadores construyen su marca personal a través de servicios web ampliamente adoptados por la comunidad académica. Estos servicios brindan por lo general algún tipo de identificador de autor, que es controlado y gestionado por la organización a cargo de cada servicio, para los que prometen persistencia y procedimientos en caso de que dicha empresa deje de brindar el servicio (y, por lo tanto, de administrar los identificadores). ORCID es un ejemplo de un conjunto de servicios similares como Google Scholar Profiles, Scopus Author ID o ResearcherID, que facilitan la integración del investigador en el ámbito académico (de ahí su alto grado de adopción). En todos estos casos, los proveedores de cada servicio deciden la manera de gestionar y exponer los datos de los usuarios. Por ejemplo, la forma en que las publicaciones son expuestas por defecto en Google Scholar Profiles está determinada por cantidad de citas, criterio definido de manera unilateral por el proveedor del servicio (Google). El uso de esta métrica, definida por Google Scholar, para ordenar las publicaciones puede no coincidir con el criterio que mejor se ajusta a las necesidades del investigador, priorizando publicaciones que no necesariamente sean las que mejor representen su perfil, ya sea porque no sea su línea de trabajo actual, porque se trate de trabajos que el autor no desea priorizar, o por el motivo que fuere.

Para muchos investigadores el servicio gratuito ofrecido por ORCID es suficiente, y las instituciones pueden contratar el servicio de ORCID para Organizaciones Científicas, lo que les permite sumar servicios a partir de los identificadores de sus investigadores<sup>3</sup>. Los servicios de ORCID para individuos e instituciones han generado un alto grado de adopción en la comunidad científica (PETRO, 2020), con más de 10 millones de perfiles hasta el mes de noviembre de 2020. ORCID se hace entonces responsable de administrar y mantener la URI y URL del perfil de cada autor. Este tercero "de confianza" se encarga de controlar los identificadores personales, y los investigadores son responsables de actualizar el contenido (CAPADISLI, 2020). Una consecuencia

---

<sup>3</sup> ORCID for Research Organizations: <https://info.orcid.org/orcid-for-research-organizations/>

involuntaria de esta adopción generalizada de ORCID en la comunicación académica es que un número creciente de instituciones y editores establecen como requerimiento<sup>4</sup> el uso de identificadores ORCID para identificar unívocamente a los actores (autores de artículos, directores proyectos de investigación, etc.), lo que a su vez fuerza a los investigadores a crear un perfil y almacenar sus datos en un espacio que no pertenece a la institución y que es controlado por una organización de la que participan empresas privadas o grupos externos con intereses diferentes. Tales iniciativas son contrarias a la participación autónoma, en el sentido de que un investigador tiene que ajustarse a los términos y políticas de terceros, externos a su propia institución.

## NGRWG

En base al escenario mencionado, uno de los principios<sup>5</sup> en los que se basa el trabajo que vienen llevando a cabo el Next Generation Repository Working Group (NGRWG), es distribuir el control de los recursos y la infraestructura que conforman los repositorios hoy en día, con el objetivo de generar redes distribuidas más sostenibles buscando disminuir, no sólo las probabilidades de errores, sino también evitar que un pequeño número de actores puedan tener demasiado control y establecer una posición cuasi monopolista. Una red social distribuida global, pensada en los términos que propone el NGRWG, requiere que cada persona pueda controlar su identidad, pudiendo vincularla a través de diversos sitios o sistemas con otros recursos o agentes (personas, organizaciones, etc.), ubicando a cada persona en una red de relaciones<sup>6</sup>. Una de las tecnologías apuntadas para llevar a cabo este desafío es precisamente Web Identity and Discovery (WebID).

---

<sup>4</sup> Requiring ORCID publication workflows open letter: <https://orcid.org/content/requiring-orcid-publication-workflows-open-letter>

<sup>5</sup> COAR Next Generation Repositories: Principles: <https://ngr.coar-repositories.org/principles>

<sup>6</sup> COAR Next Generation Repositories: Identification of Users: <https://ngr.coar-repositories.org/behaviour/identification-of-users/>

WebID (SAMBRA *et al.*, 2014) es una propuesta de la W3C realizada en 2014, cuya versión final aún no ha sido publicada oficialmente. WebID ofrece un mecanismo de identificación universal simple, distribuido, extendido que mejora el control sobre la forma en que cada persona puede identificarse para permitir un control de acceso detallado a su información en la web. Se trata de un mecanismo distribuido, ya que al estar formado por una URI puede ser referenciado desde cualquier lugar de la web. Además, WebID es extensible, ya que puede ser utilizado como base para el uso de otros protocolos como por ejemplo un mecanismo de autenticación y control de acceso seguro como WebID-OIDC, WebID-TLS o WebID-TLS+Delegation (MAININI & LAUBE-ROSENPFLANZER, 2016). Si bien existen entidades como ORCID que ya presentan esta clase de servicios, donde el autor puede publicar sus datos y asociarlos a uno o más identificadores, WebID brinda la posibilidad de definir un modelo flexible, que se ajuste a las necesidades del usuario, ya que utiliza ontologías para definirlo.

## WebID y WebID-Profile

Como se indicó previamente, WebID es un mecanismo extensible, descentralizado y distribuido que permite identificar a un agente (persona, organización, grupo, etc.) a través de una URI dentro de un dominio que idealmente sea propiedad del agente a identificar. Esta URI referencia a un perfil que puede ser interpretado por cualquier usuario web, representado en un archivo bajo el control del agente, llamado WebID-Profile, donde se pueden especificar cuestiones como qué publicaciones posee una persona, a quiénes conoce, cuáles son sus intereses, en qué proyectos trabaja, cómo se conforma su *curriculum vitae*, qué fotos posee, cuáles son sus claves públicas, si utiliza otros identificadores, etc.

El objetivo de WebID es que los agentes creen sus propias identidades, relacionando un identificador único a un perfil que pueda ser interpretado por cualquier usuario web. Para realizar esto, WebID hace uso de serialización en

formato RDF. Además, un perfil puede contener toda la información necesaria para construir relaciones con otros recursos de la web que pueden ser utilizados por distintos servicios web.

WebID Profile<sup>7</sup> es un documento RDF que describe al agente a través del uso de diversos vocabularios como FOAF<sup>8</sup> y OWL<sup>9</sup>. Este documento puede almacenarse en el propio servidor web del agente, y su acceso puede estar parcial o totalmente restringido a terceros mediante el uso de controles de acceso, para preservar la privacidad del sujeto del documento de perfil. Dado que el WebID Profile es un documento RDF, su dueño puede utilizar diversas ontologías para modelar la información que crea conveniente y así poder extender su perfil, utilizando los principios propuestos por Linked Data (BERNERS-LEE, 2009), una tecnología que busca dotar de semántica a la web modelando los recursos representados en ella y estableciendo relaciones entre ellos.

Una de las dificultades que se pueden encontrar en la implementación de este protocolo, es la necesidad de disponer de un espacio donde alojar el WebID Profile. Como dijimos anteriormente, es deseable que este documento resida en un servidor propiedad del dueño del WebID Profile, pero la realidad indica que no todas las personas tienen esa posibilidad. Este impedimento hace necesario realizar un esfuerzo por parte de las instituciones para poder brindar a sus investigadores un espacio que no solo permita almacenar estos perfiles, sino que garantice su disponibilidad a largo plazo, en caso de que no dispongan de un servidor de tales características.

A continuación se puede observar el EJEMPLO 1, que representa un WebID Profile donde se observa cómo se utilizan los vocabularios FOAF y RDF para describir a una persona.

---

<sup>7</sup> WebID Profile or Profile Document  
[https://www.w3.org/2005/Incubator/webid/spec/identity/#dfn-profile\\_document](https://www.w3.org/2005/Incubator/webid/spec/identity/#dfn-profile_document)

<sup>8</sup> FOAF Vocabulary Specification 0.99 <http://xmlns.com/foaf/spec/>

<sup>9</sup> OWL - Semantic Web Standards <https://www.w3.org/OWL/>

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

<> a foaf:PersonalProfileDocument;
  foaf:maker <#me>;
  foaf:primaryTopic <#me>.

<#me> a foaf:Person;
  foaf:name "Juan";
  foaf:family_name "Gonzalez";
  foaf:mbox <user@example.com>.
```

EJEMPLO 1. Un WebID Profile expresado en N3

En dicho documento es posible notar la presencia de dos sujetos: una URI vacía <>, que significa “este documento”, y el recurso <#me>, que hace referencia a la persona que describe el WebID Profile. Del documento (<>) podemos ver cómo se desprenden dos predicados, foaf:maker y foaf:primaryTopic cuyos objetos hacen referencia al recurso <#me> indicando no solamente que este es quien creó el documento sino que también es el objetivo del documento describir a esa persona. Del recurso <#me> podemos saber que es del tipo foaf:Person a través de la propiedad a, (que es un alias para rdf:type). De esta persona sabemos que su nombre es "Juan", de apellido "Gonzalez" y cuyo mail es <user@example.com>, a través de las propiedades name, family\_name y mbox del vocabulario foaf. Dicho más simple, el ejemplo anterior muestra cómo se puede relacionar un WebID Profile, dedicado a describir a la persona “Juan Gonzalez”, indicando que este es su creador (*maker*).

## Extensión del WebID Profile

Una de las ventajas del uso de WebID es la facilidad para extenderlo, incorporando nueva información y nuevas relaciones. Para lograr esto solo es necesario expresar la nueva información a través de otros vocabularios. Por

ejemplo, en el EJEMPLO 2 es posible ver cómo a través de los vocabulario SIOC y DC, se especifica que la persona “Juan Gonzalez”, descrita en el documento, tiene un posteo en su blog, especificando el título del posteo (My first post) y su contenido (Lorem ipsum).

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix sioc: <http://rdfs.org/sioc/ns#>.
@prefix dc: <http://purl.org/dc/elements/1.1/>.

<> a foaf:PersonalProfileDocument;
    foaf:maker <#me>;
    foaf:primaryTopic <#me>.

<#me> a foaf:Person;
    foaf:name "Juan";
    foaf:family_name "Gonzalez";
    foaf:mbox <user@example.com>.

<https://juan.example.com/blog/posts/1> a sioc:Post;
    sioc:has_creator <#me>;
    dc:title "My first post";
    sioc:content "Lorem ipsum".

<https://myproject.org/> a doap:Project;
    doap:name "My Project";
    doap:developer <#me>.
```

EJEMPLO 2. WebID Profile que describe un posteo en un blog y un proyecto utilizando el vocabulario SIOC y DOAP

En el ejemplo anterior se utiliza también el vocabulario DOAP, que permite describir los proyectos en los que está involucrada una persona, con la posibilidad de especificar el rol que desempeña en él. Gracias a las ventajas que trae el uso de Linked Data, cualquier servicio podría referenciar la URI del proyecto (en este caso [<https://myproject.org/>](https://myproject.org/)) para describir más información.

## Uso de múltiples identificadores en WebID Profile

Cuando un usuario crea una cuenta en algún servicio o red social, se le asigna un identificador que sirve en el contexto de dicha aplicación, por lo que es normal que un investigador disponga de muchos identificadores (por ejemplo, ORCID, Scopus Author ID, WoS ResearcherID, etc.). WebID permite asociar a un perfil otros identificadores que a su vez referencian a otros perfiles y que pueden aportar más información sobre el autor. Una buena práctica es editar los perfiles definidos en otras aplicaciones para que hagan referencia al WebID que el investigador puede controlar. Por ejemplo, a través de la interfaz web de ORCID, un investigador puede modificar su perfil y relacionarlo con otras URI. De esta forma se crea una relación semántica entre el perfil ORCID y el WebID a través de una propiedad owl:sameAs lo que permite que otros programas o agentes informáticos puedan acceder información que no está publicada en ORCID, ya sea porque el modelo no permite incorporar ese dato o simplemente porque no se encuentra disponible.

ORCID ofrece una API a través de la cual es posible acceder a una vista de un perfil en formato RDF. En el EJEMPLO 3 es posible ver como en [https://pub.orcid.org/experimental\\_rdf\\_v1/0000-0002-0880-9125](https://pub.orcid.org/experimental_rdf_v1/0000-0002-0880-9125) el autor Sarven Capadisli referencia WebID Profile, del que no solamente es dueño sino que también tiene autoridad a nivel de DNS sobre el dominio csarven.ca, donde tiene publicado dicho documento. Para establecer esta relación entre ORCID y su propio WebID profile se utilizó la propiedad owl:sameAs .

```

@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .

<https://orcid.org/0000-0002-0880-9125>
  a foaf:Person;
  foaf:name "Sarven";
  foaf:family_name "Capadisli";
  owl:sameAs <http://csarven.ca/#i> .

```

EJEMPLO 3. Uso de owl:sameAs en RDF del perfil de investigador Sarven Capadisli en ORCID

Es así como las aplicaciones que acceden a la información que brinda ORCID, podrán acceder a través de la propiedad owl:sameAs a la información que el investigador publica en su propio WebID Profile, como puede ser una foto o imagen del investigador, dato que no es posible cargar en ORCID.

En la siguiente sección daremos un ejemplo de un caso concreto en el que un repositorio institucional puede incorporar la información publicada por un autor en un WebID Profile que esté asociado a múltiples identificadores.

## Uso de WebID en un repositorio institucional

Anteriormente se mencionó cómo el desarrollo de la marca personal puede mejorar la reputación de un investigador, a través del control de la información publicada en la web. En su artículo, CLIMENT RODRÍGUEZ y NAVARRO-ABAL (2017) afirman que la marca de una institución, y su reputación, constituyen la suma de las reputaciones de las marcas de sus miembros más relevantes. A raíz de esto GONZÁLEZ SOLAR (2018) comenta que una institución “no debe buscar potenciar la autopromoción de los investigadores sino favorecer que estos puedan gestionar de forma racional su marca, dotándolos de las herramientas y conocimientos necesarios y evitando la pérdida de tiempo y esfuerzo en actividades que se pueden realizar colectivamente”.

Una propuesta de muchos repositorios institucionales que resuelven este tipo de cuestiones se basa en incorporar información de sus investigadores a partir de una fuente de datos normalizada, como puede ser un sistema de autoridades propio o de terceros. De esta forma el repositorio establece relaciones entre sus recursos con entidades normalizadas como autores o instituciones. Es así que proyectos como los del software DSpace, utilizado como base para implementar repositorios digitales, permite desarrollar un conector personalizado para consumir datos de otras fuentes.

La gestión de autoridades en un repositorio institucional puede realizarse desde el mismo software del repositorio, como por ejemplo en la nueva versión de DSpace 7, donde se utiliza un nuevo concepto llamado entidades configurables (DROOGMANS, 2019), que permite expandir el modelo, incorporando entidades que suelen utilizar sistemas CRIS, como personas, instituciones, proyectos, etc. Otra forma de gestionar autoridades en un repositorio institucional es a partir de un software dedicado exclusivamente a la gestión de estos y que comparta sus datos con el software del repositorio. En DE ALBUQUERQUE (2018) se describe este último modelo, basado en Drupal 7 como sistema de gestión de autoridades, y DSpace 5 como software del repositorio. En dicho trabajo, las entidades se relacionan a través de identificadores definidos internamente, con los recursos que administra el repositorio.

Ante cualquiera de estos escenarios, el repositorio podría armar perfiles de autores que fomenten la visibilidad de los investigadores, tratándolos como una entidad referenciable en la web, permitiendo agrupar información proveniente de diversas fuentes de datos. Con los datos agrupados en estos perfiles de datos es posible generar información y ofrecer estadísticas de uso basado en las publicaciones gestionadas dentro del repositorio, o inferir redes de coautoría a partir de la relación que un investigador tiene con su producción, que a su vez se relaciona con otros autores. Las fuentes de datos a utilizar para generar estos perfiles podrán basarse en, no sólo los datos del repositorio, sino también en lo que un investigador publique en su WebID Profile, junto con las

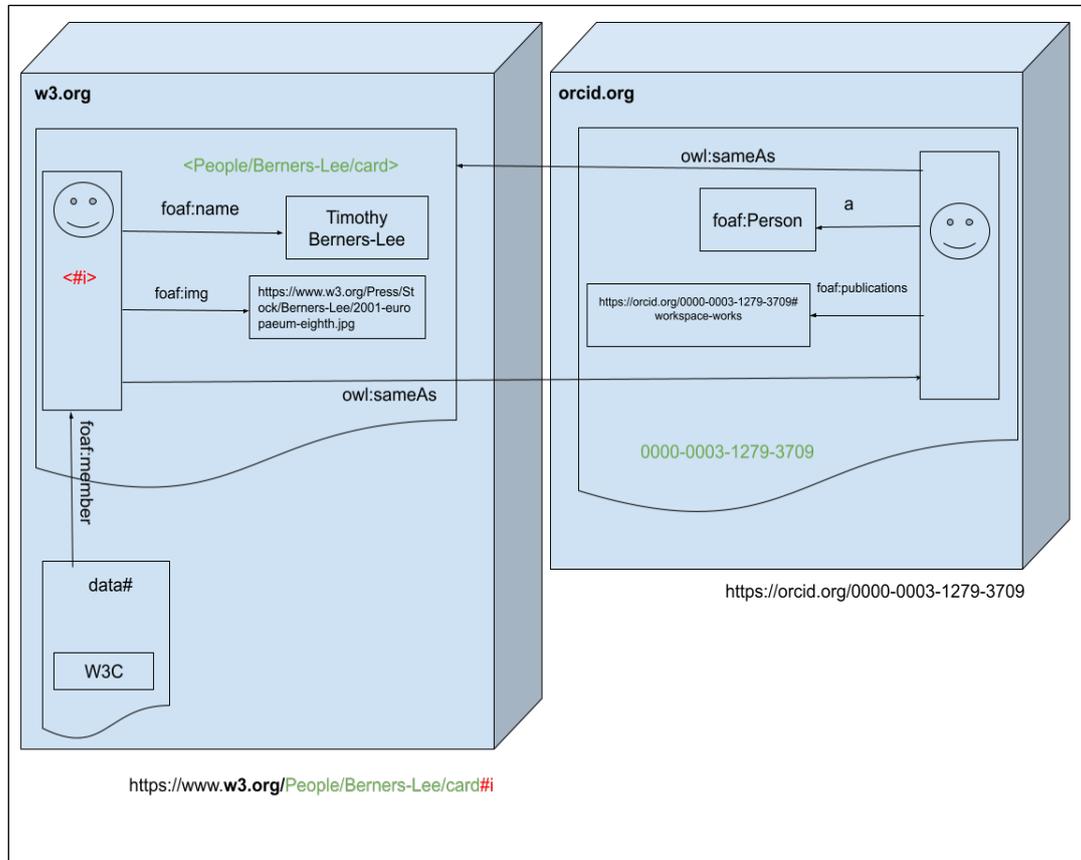
referencias que se hagan en otros perfiles, como ORCID o Google Scholar Profile.

La forma en que se puede incorporar el uso de WebID en el contexto del repositorio puede realizarse almacenando el WebID del investigador como clave de la autoridad que lo representa. Para ello, antes de implementar cualquier cambio en el software del repositorio, es importante aclarar que su esquema de metadatos debe poder referenciar al investigador como una entidad normalizada. Esto significa que los autores deberán ser entidades controladas, deberán realizarse esfuerzos para unificar autores, normalizar sus nombres, actualizar sus filiaciones institucionales y realizar cualquier otra tarea necesaria para asegurar la calidad de este metadato. Dicho de otro modo, los repositorios que representen los metadatos “autor” como texto simple (*string*) entre los metadatos de los objetos digitales, no podrán incorporar un servicio de calidad basado en WebID profile. También existe la alternativa intermedia, en la que algunos autores de objetos digitales se asocian a una entidad normalizada (autores propios de la institución) y otros autores son representados como un texto y no se vinculan a entidades normalizadas (autores externos de la institución); en este modelo, sólo los autores vinculados a entidades normalizadas del repositorio podrán utilizar el servicio de WebID profile.

Una vez almacenado el WebID del investigador, el repositorio podría obtener más datos de este, y ofrecer nuevos servicios, como la implementación de perfiles de autor a partir de los datos ya normalizados por el autor, o la inclusión de la URI del investigador junto a los metadatos de los objetos digitales que expone mediante las diferentes interfaces (HTML, OAI-PMH, REST-API, etc.).

En el EJEMPLO 4 se ilustra parte de la información que se publica en el WebID profile del investigador Tim Berners Lee, identificado con la URI <https://www.w3.org/People/Berners-Lee/card>.

El recurso que hace referencia al investigador es <https://www.w3.org/People/Berners-Lee/card#i> y de allí se puede obtener el nombre, una imagen y saber que es miembro del W3C.



EJEMPLO 4. Parte de la información expuesta en el WebID Profile de Tim Berners Lee

Como los datos del EJEMPLO 4 se modelan a través de la ontología FOAF, es a través de la propiedad `foaf:img` que un repositorio podría acceder a la foto publicada por el autor, evitando almacenar y mantener imágenes que podrían no ser el principal interés para un repositorio. La información recuperada del WebID Profile del autor, sumada a la que dispone el repositorio, permitiría no sólo implementar perfiles de usuarios más completos sino aumentar la visibilidad del autor en otros sistemas que también hagan referencia al mismo recurso.

En el mismo ejemplo también se puede observar cómo la propiedad definida como owl:sameAs en el perfil alojado en w3.org referencia al perfil del autor en orcid.org. A partir de esta relación es posible obtener, a partir de la propiedad foaf:publication, la URL con el listado de publicaciones que el autor cargó en ORCID y utilizar esa información para enriquecer el perfil de usuario creado en el repositorio.

## Conclusión

En este trabajo se analizó la importancia que tiene para un investigador controlar su identidad en la web y cómo esta puede ayudar al desarrollo de una marca personal. En este desafío de difundir la actividad académica de un investigador, se ha involucrado el trabajo de los repositorios institucionales y cómo se puede utilizar WebID como complemento a los servicios más utilizados hoy día, como ORCID y Google Scholar Profiles. Es importante resaltar que el uso de WebID no busca reemplazar a los identificadores controlados por terceros, sino más bien darle al autor una herramienta que le otorgue mayor control sobre su identidad en la web.

Si bien WebID es una propuesta que data del año 2014 y su versión final aún no ha sido publicada oficialmente, el hecho de que sea una tecnología considerada por el NGRWG y de que existan implementaciones concretas, hacen de WebID una tecnología a considerar para ser integrada en el desarrollo de servicios pensados desde repositorios institucionales. Su uso puede significar no solo un acercamiento hacia la descentralización de la arquitectura de los repositorios, sino también un espacio para que los investigadores puedan construir su propia marca personal, facilitando tanto el desarrollo de la reputación del investigador como la de las instituciones con las cuales se relaciona, entendiendo a la reputación de la institución como una construcción de la suma de las reputaciones de sus investigadores.

La incorporación de WebID en una institución puede ser un camino que ayude al trabajo de normalización de la información que se tiene sobre los investigadores de una forma descentralizada, cediendo parte de este trabajo al mismo investigador. De esta forma sería posible reutilizar el WebID en distintas fuentes de datos, no solamente dentro de la institución sino con otros sistemas.

Es importante resaltar que el WebID debe ser gestionado por el investigador y el lugar donde se encuentre alojado el WebID Profile es su responsabilidad. Sin embargo, en caso de que el investigador no posea acceso a un servidor propio dedicado a este fin, podría ser la misma institución la que brinde este espacio como un servicio no sólo para construir la marca personal del investigador sino también para tener un mayor control sobre la reputación de la institución en la web. Es evidente que esto agrega una carga adicional tanto a la infraestructura informática de la institución como al personal técnico a cargo de mantener dicha infraestructura, y por lo tanto cada institución debe considerar de manera cuidadosa las ventajas y los costos de incorporar este tipo de servicios para sus miembros.

## Bibliografía

- BERNERS-LEE, T. (2009, junio 18). Linked Data—Design Issues. <https://www.w3.org/DesignIssues/LinkedData.html>
- BROWN, R. (2010). Reputation management. *Business Information Review*, 27(1), 56-64. <https://doi.org/10.1177/0266382109357390>
- CAPADISLI, S. (2020). Linked Research on the Decentralised Web. <http://hdl.handle.net/20.500.11811/8352>
- CLIMENT-RODRÍGUEZ, J. A., & NAVARRO-ABAL, Y. (2017, febrero). Branding y reputación: Pilares básicos de la visibilidad online del profesor de educación superior. <http://ref.scielo.org/h2nc8c>

- DE ALBUQUERQUE, P. C. (2018). Soporte de vocabularios controlados y autoridades en repositorios digitales [Tesis, Universidad Nacional de La Plata]. <http://sedici.unlp.edu.ar/handle/10915/69754>
- DROOGMANS, L. (2019, noviembre 26). DSpace 7—The Power of Configurable Entities. The 14th International Conference on Open Repositories (or2019), Hamburg, Germany. <https://doi.org/10.5281/zenodo.3554102>
- GONZÁLEZ-SOLAR, L. (2018). Marca personal en entornos académicos: Una perspectiva institucional. *Anales de Documentación*, 21(2), Article 2. <https://doi.org/10.6018/analesdoc.21.2.328561>
- MAININI, P., & LAUBE-ROSENPFANZER, A. (2016). Access Control in Linked Data Using WebID. arXiv:1611.03019 [cs]. <http://arxiv.org/abs/1611.03019>
- PETRO, J. (2020, octubre 20). 10M ORCID iDs! 10M ORCID IDs! <https://info.orcid.org/10m-orcid-ids/>
- SAMBRA, A., STORY, H., & BERNERS-LEE, T. (2014, mayo 28). WebID 1.0. <https://dvcs.w3.org/hg/WebID/raw-file/tip/spec/identity-respec.html>

# Conectando fortalezas y capacidades institucionales: automatización de la integración VIVO-Pure-DSpace

## **MALGORZATA LISOWSKA NAVARRO**

Universidad del Rosario (Colombia)

[margarita.lisowska@urosario.edu.co](mailto:margarita.lisowska@urosario.edu.co)

## **HUMBERTO BLANCO CASTILLO**

Universidad del Rosario (Colombia)

[humberto.blanco@urosario.edu.co](mailto:humberto.blanco@urosario.edu.co)

## **ÁNGELICA CHARTANO HERNÁNDEZ**

Universidad del Rosario (Colombia)

[angelica.chartano@urosario.edu.co](mailto:angelica.chartano@urosario.edu.co)

## **JUAN FELIPE LÓPEZ**

Universidad del Rosario (Colombia)

[juan.lopez@urosario.edu.co](mailto:juan.lopez@urosario.edu.co)

## **RESUMEN**

El HUB-UR: Services and Experts Finder, portal de fortalezas y capacidades de la Universidad del Rosario, implementado en el año 2018 por el Centro de Recursos para el Aprendizaje y la Investigación (CRAI), recibe anualmente más de 110.000 visitas de cien países. El HUB-UR, basado en el software de web semántica y de código abierto VIVO permite, a través de búsquedas y filtros, descubrir las fortalezas y capacidades en investigación, docencia y extensión de la universidad, aumentando así su reconocimiento a nivel nacional e internacional. En

sus tres años de funcionamiento, el reto continúa siendo obtener la información actualizada y automatizada de otras fuentes institucionales como el Repositorio Institucional E-docUR (DSpace) y el sistema Pure, software de la casa editorial Elsevier, a través del cual se gestiona la investigación en la universidad. Dado que no existe abundante literatura acerca de la interoperabilidad de VIVO con estas plataformas, la presente ponencia pretende ilustrar la forma a través de la cual se ha desarrollado la integración entre VIVO, Pure y DSpace. Nuestro interés se centra en compartir la experiencia y los retos que se presentaron en el proceso de automatización y que puedan servir a otras instituciones interesadas en este proceso de integración.

#### **PALABRAS CLAVE**

Interoperabilidad; repositorios académicos; visibilidad académica; VIVO.

*Interoperability; academic repositories; academic visibility; DSpace; Pure.*

El HUB-UR: Services and Experts Finder, basado en el software VIVO aprovecha las ventajas de datos enlazados abiertos (Linked Open Data)<sup>1</sup>, para relacionar y vincular la información de perfiles individuales, institucionales, proyectos, productos de investigación y eventos, entre otros (CONFLUENCE, 2022), proporcionando así resultados hipervinculados en forma de servicios para diferentes públicos objetivo. Por ejemplo, información para estudiantes en proceso de grado a través del servicio “*Find a supervisor*” o información para pares de investigación a través del servicio “*Find a research partner*” .

---

<sup>1</sup> Linked Data se refiere a datos enlazados a través de web semántica, la manera de describir y representar esta información se realiza mediante el estándar RDF (Resource Description Framework).

Para visibilizar las capacidades y fortalezas tanto individuales como institucionales, fue necesario integrar diferentes fuentes de información institucional (ver GRÁFICO 1), esto con el fin de asegurar su calidad, actualización y confiabilidad.

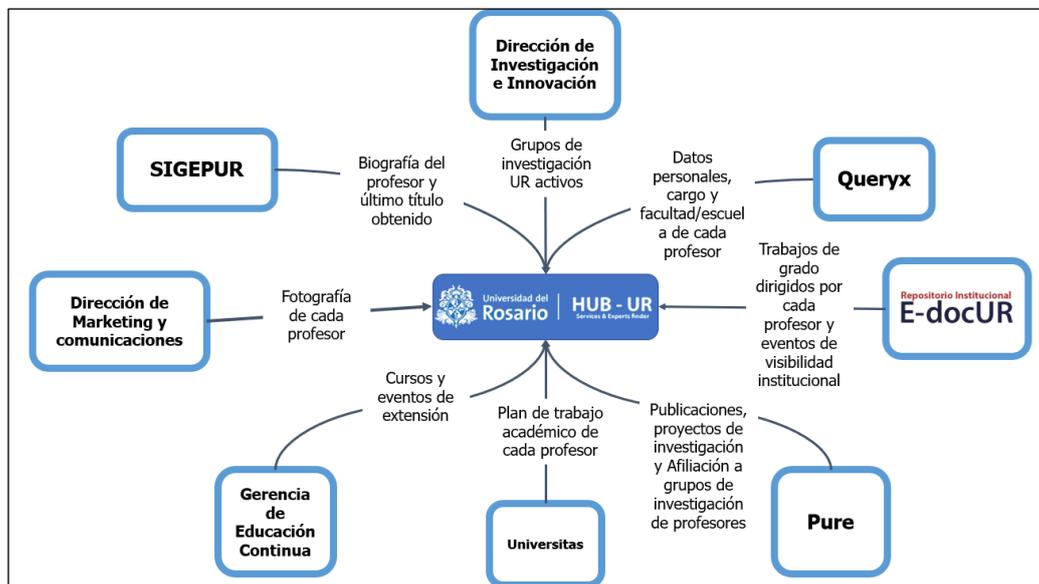


GRÁFICO 1. Diagrama de sistemas. Fuente: HUB-UR

Esta integración ha constituido un reto, pues los sistemas de información fuente han sido desarrollados en diferentes momentos de la historia de la universidad, lo que conlleva al registro de información de diferentes maneras; por esta razón, algunos no cuentan con interfaces que permitan la interoperabilidad, sino que, de forma manual, se realiza la extracción, transformación y carga de información en VIVO.

Actualmente, se ha logrado la automatización con los sistemas de gestión de investigación Pure de Elsevier y DSpace, los cuales proveen principalmente información sobre producción académica y actividad científica de la universidad. El Sistema de Gestión de la Investigación - Pure<sup>2</sup> es la plataforma

<sup>2</sup> Pure: <https://www.elsevier.com/solutions/pure>

utilizada para la gestión de la investigación en la UR, que contiene información sobre proyectos y resultados de investigación; actualmente el software Pure se encuentra en la versión 5.21. Por otra parte, el Repositorio Institucional E-docUR, es la plataforma institucional que almacena, preserva y difunde la producción institucional en su función docente, investigativa y de extensión, está implementado en el software de código abierto DSpace en su versión 6.3.

## Proceso de automatización de la integración VIVO-DSpace-Pure

### 1. Integración del sistema de gestión de investigación UR-Pure-VIVO

El proceso de extracción y carga de información en Pure se ha realizado de acuerdo con el siguiente diagrama (DIAGRAMA 1).

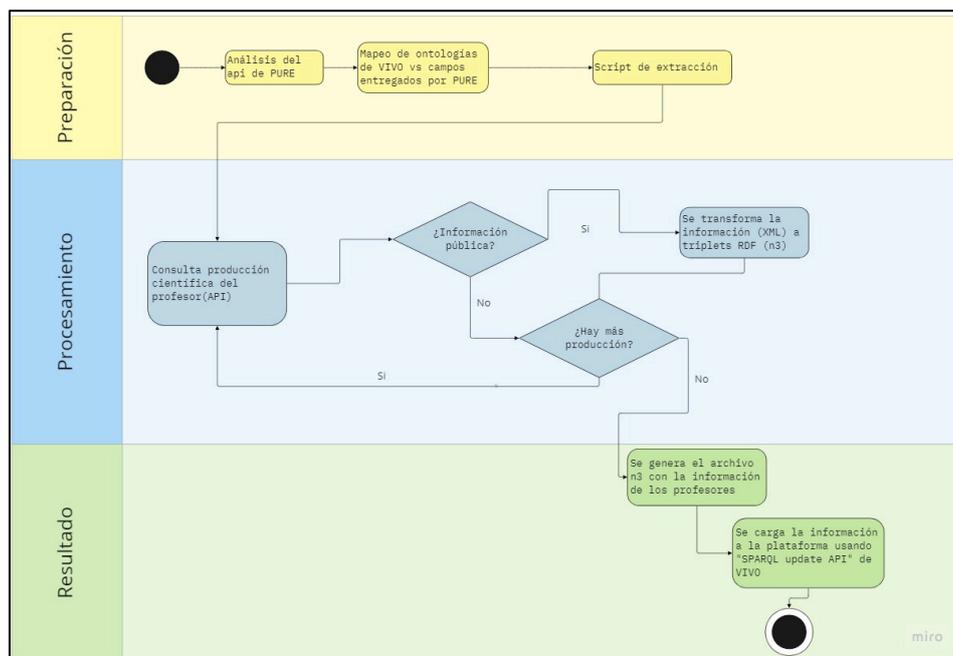


DIAGRAMA 1. Etapas de integración VIVO-Pure

La fase de preparación para implementar la solución partió del análisis de la API de Pure (ELSEVIER, 2022), para ello fue necesario solicitar al proveedor del servicio (Elsevier) una API KEY, realizar su conexión y empezar la exploración de la API. En esta fase, fue necesario realizar ejercicios para validar la forma y contenidos en que la API entregaba la información para ser consumidos, además de los requerimientos para su conexión.

Posteriormente, fue necesario realizar un mapeo de la información disponible en Pure para determinar su correspondiente ontología en VIVO, requerida para almacenar la información que sería extraída. A continuación, se presenta un ejemplo del mapeo realizado (TABLA 1):

<b>ATRIBUTO</b>	<b>ONTOLOGÍA</b>
Título	<i>22-rdf-schema:label</i>
Keywords	<i>vivo:freetextkeyword</i>
Tipo	<i>22-rdf-syntax-ns:type</i>
Número de páginas	<i>bibo:numpages</i>
Abstract	<i>Core:overview</i>
autores	<i>Core:relatedBy</i>
Volumen	<i>bibo:volume</i>
Edición	<i>bibo:issue</i>
ISSN	<i>bibo:issn</i>
DOI	<i>Vcard:hasURL</i>
Idioma	<i>vcard:language</i>
Journal	<i>vlocal:academicarticlejournal</i>
Fecha de publicación	<i>vivo:datetimevalue</i>
Página de inicio	<i>bibo:pagestart</i>

ATRIBUTO	ONTOLOGÍA
ISBN	<i>vlocal:isbn</i>
Evento	<i>vlocal:conferencePaperEvent</i>
Número de patente	<i>core:patentNumber</i>

TABLA 1. Mapeo de atributos Pure-Ontología VIVO

La fase que contempla la transformación de los atributos a triples al inicio del proyecto, fue realizada empleando el software Karma, el cual permite generar los archivos con la información procesada para ser cargada a VIVO (CONLON, 2021), sin embargo, este proceso fue optimizado posteriormente con la creación de un *script* que automatizó el proceso de extracción de información del API de PURE y su transformación en tripletas RDF, formato n3 (W3C, 2011).

La fase de procesamiento que inicia desde la extracción con la consulta a la API de Pure de los datos personales e identifica de la producción de cada uno de los de los profesores activos en el HUB-UR, esta consulta se realiza con base en el documento de identificación (ID) de cada profesor.

El *script* se encarga de mapear la respuesta entregada por el API de Pure a las ontologías correspondientes de VIVO, esto garantiza que los productos de investigación y sus metadatos tengan las clases y propiedades de datos correspondientes en el HUB-UR. Al finalizar este procesamiento su resultado es un archivo generado por el *script* en formato n3 con la información recopilada de los productos, los cuales posteriormente son cargados a VIVO utilizando el "SPARQL update API", que brinda capacidades para consultar patrones gráficos requeridos y opcionales junto con sus conjunciones y disyunciones (W3C, 2022).

## 2. Integración del repositorio institucional E-docUR (DSpace)-VIVO

Como se ha indicado, el Repositorio institucional E-docUR (DSpace) alberga dentro de sus contenidos los trabajos de grado y tesis que han sido dirigidas por los profesores UR así como los productos asociados a su participación en conferencias y eventos, estos contenidos al considerarlos parte de las fortalezas y capacidades de los profesores son integrados en el HUB-UR.

La extracción de la información que proviene de DSpace es realizada directamente desde la base de datos del repositorio, y se extrae por medio de *scripts* de acuerdo con el contenido a extraer; uno para las tesis y trabajos de grado dirigidos, otro para las conferencias y otro para los eventos.

Para la extracción de tesis y trabajos de grado la consulta es realizada a través del documento de identidad del profesor (ID) que para DSpace corresponde al valor "Authority" de la tabla *metadatavalue*. El siguiente diagrama muestra las diferentes etapas en el proceso (DIAGRAMA 2).

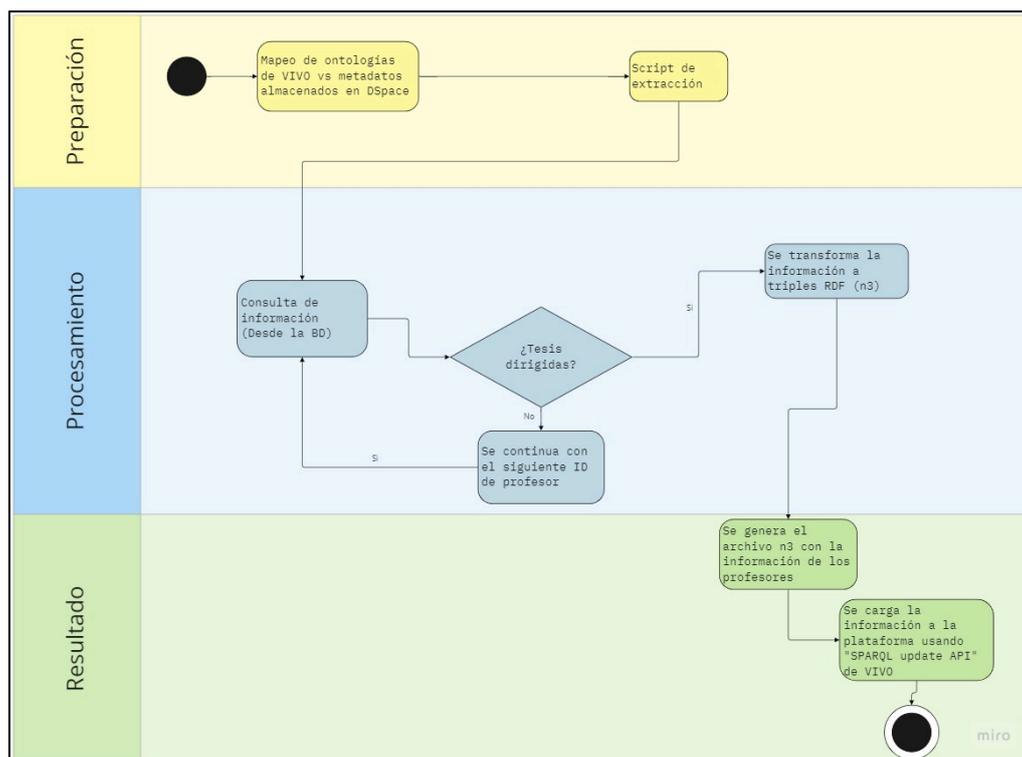


DIAGRAMA 2. Etapas de integración VIVO-DSpace (tesis y trabajos de grado)

Para el caso de las tesis y trabajos de grado el mapeo de los atributos en DSpace y su correspondiente ontología en VIVO se realiza sobre los siguientes campos (TABLA 2):

ATRIBUTO	ONTOLOGÍA
Titulo	<i>22-rdf-schema:label</i>
Keywords	<i>vivo:freetextkeyword</i>
Tipo	<i>22-rdf-syntax-ns:type</i>
Abstract	<i>Core:overview</i>
autores	<i>Core:relatedBy</i>
Idioma	<i>vcard:language</i>
Fecha de publicación	<i>vivo:datetimevalue</i>

TABLA 2. Ejemplo de mapeo atributo DSpace -Ontología VIVO (tesis y trabajos de grado)

Una vez consolidada y procesada la información se envía a través de la API de actualización de VIVO donde el resultado es el archivo generado por el *script* en formato n3, el cual es cargado a VIVO utilizando el “SPARQL update API”.

Para la extracción de eventos y conferencias es necesario realizar una identificación previa de las colecciones que alojan esos productos en DSpace, a partir de esta identificación el *script* realiza la consulta de los participantes (autores) en los eventos y conferencias y valida su existencia en un perfil de VIVO. Al ser identificado, el *script* genera el triples con los datos del evento o conferencia (DIAGRAMA 3).

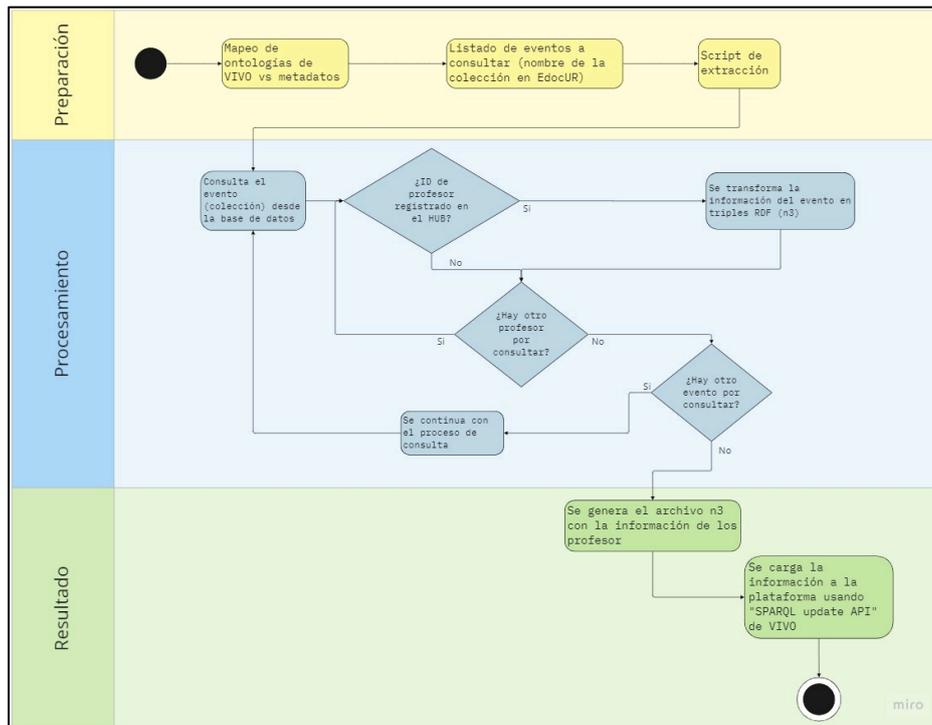


DIAGRAMA 3. Etapas de integración VIVO-DSpace (eventos y conferencia)

Para los eventos y conferencias los atributos en DSPACE y su correspondiente ontología en VIVO es la siguiente tabla:

ATRIBUTO	ONTOLOGÍA
Titulo	<i>22-rdf-schema:label</i>
Keywords	<i>vivo:freetextkeyword</i>
Tipo	<i>22-rdf-syntax-ns:type</i>
Abstract	<i>Core:overview</i>
Temáticas	<i>vivo:hassubjectarea</i>
Participantes	<i>vlocal:participants</i>
Descripción	<i>vivo:description</i>
Fecha	<i>vivo:datetimeinterval</i>

TABLA 3. Ejemplo de mapeo atributo DSpace -Ontología VIVO (eventos y conferencias)

## Retos identificados en el proceso

El proceso de extracción carga y actualización de la información ha supuesto retos que se han superado gracias a los aprendizajes obtenidos a medida que el proyecto ha ido madurando. Algunos de ellos se presentan a continuación:

- Traducir la información de Pure y DSpace, a un lenguaje comprensible para VIVO fue un proceso complejo. Esto debido a que este último utiliza ontologías que requieren archivos en formato n3. Para lograr esto ha sido necesario comprender las consultas SPARQL, con el fin de identificar las clases, relaciones y ontologías que se deben asignar a cada tipología extraída de Pure y DSpace.
- Para garantizar la funcionalidad del procedimiento ha sido necesario estar alineados a los cambios generados en las nuevas versiones de la API de Pure, lo que implica un monitoreo permanente y aplicar cambios periódicos en los *scripts*.
- El trabajo interdisciplinar con otras áreas exige una comunicación constante y el diseño de flujos de trabajo que permita dar respuesta oportuna ante los cambios en las plataformas.

## Impactos obtenidos

La implementación de estos procedimientos generó resultados significativos, dentro de los que se encuentran:

- Reducción en los tiempos de gestión: al automatizar los procesos de extracción de información, se logró reducir los tiempos de gestión que implicaban revisiones y ajustes manuales de la información, especialmente en el momento de actualización de contenidos.
- Información actualizada y fiable: durante el proceso de extracción de información de sistemas fuente, ha permitido identificar información incompleta, duplicada o desactualizada, que ha implicado el trabajo conjunto de depuración y corrección de la información con las áreas responsables con lo cual se ha logrado unificar la información

institucional y mejorar la calidad de la información disponible en los sistemas fuente.

- Reducción en frecuencia de actualización de productos: al automatizarse la extracción de información a través de los *scripts* la actualización de productos en VIVO pasó de ser semestral a trimestral lo que permite una mayor correspondencia en los productos asociados a cada profesor en las diferentes plataformas.

Se evidencia mayor concordancia institucional sobre la información visible relacionada con la producción académica y de investigación.

## Mejoras y trabajo futuro

Algunas de las mejoras que se han contemplado desde el aspecto tecnológico para el HUB-UR: Services and Experts Finder son:

- Actualmente se trabaja en la mejora de los *scripts* para que la sincronización de la información proveniente de Pure y DSpace se realice de forma inmediata.
- Se ha evaluado la creación de un nuevo *script* que permita identificar y eliminar de forma automática los registros que dejan de existir en Pure, lo cual obedece a procesos como depuraciones, correcciones o a la salida de profesores de la planta.
- Se evalúa la integración con nuevas fuentes de información institucionales que pueden proveer contenidos para VIVO.
- La interoperabilidad entre las plataformas VIVO-DSpace-Pure, permitirá en un futuro la visualización de objetos digitales complejos en el HUB-UR que faciliten la identificación y vinculación de diferentes productos vinculados al mismo proceso de investigación.

## Bibliografía

- CONFLUENCE. (2022). VIVO. VIVO 1.12.x.  
<https://wiki.lyrasis.org/display/VIVODOC112x/SPARQL+Update+API>
- CONLON, M. (2021). Karma for Data Ingestion—VIVO - LYRASIS Wiki [Wiki]. Confluence.  
<https://wiki.lyrasis.org/display/VIVO/Karma+for+Data+Ingestion>
- ELSEVIER. (2022). PURE API - 5.21. Pure API. [https://purehost.bath.ac.uk/ws/api/521/api-docs/documentation/Content/Topics/Web\\_Services\\_Intro.htm](https://purehost.bath.ac.uk/ws/api/521/api-docs/documentation/Content/Topics/Web_Services_Intro.htm)
- W3C. (2011). Notation3 (N3): A readable RDF syntax  
<https://www.w3.org/TeamSubmission/n3/>
- W3C. (2022). SPARQL Query Language for RDF. <https://www.w3.org/TR/rdf-sparql-query/>

BIREDIAL  
ISTEC



---

Pósters

---

# Caracterización de las revistas de acceso abierto en la Universidad Nacional, Costa Rica, 2022

**MARÍA AMALIA PENABAD CAMACHO**

Universidad Nacional (Costa Rica)

**LIANA PENABAD CAMACHO**

Universidad Nacional (Costa Rica)

## RESUMEN

El póster corresponde al eje temático «Comunicación académica, científica y cultural en abierto». Ante una realidad regional latinoamericana con vocación de comunicación del conocimiento en abierto, es necesario conocer cuál es el estatus de los esfuerzos que realizan las revistas para consolidar su posición individual y la institucional al respecto. El objetivo propuesto es caracterizar las revistas de la Universidad Nacional, Costa Rica (UNA) y mediante las variables identificadas, dimensionar el vínculo con el Acceso Abierto (AA) y cuáles retos tienen en su futuro cercano. Metodológicamente, se recurrió a revisar los sitios web de las 27 revistas en el Portal de Revistas Académicas de la UNA y a analizar la información disponible para identificar variables comunes; específicamente, se exploraron: licenciamiento, políticas editoriales, norma de citación, arbitraje, formatos de publicación, periodicidad, tipologías documentales, redes sociales, estadísticas, servicios de valor agregado e indexaciones. Con esta información se pudo confirmar la adopción de acciones orientadas hacia mejoras de la calidad de la gestión editorial y la aparición de acciones concretas relacionadas con la consolidación

del AA. Se identificaron retos relacionados con el uso de redes sociales para ampliar la divulgación del conocimiento y de métricas y estadísticas para conocer su alcance e impacto y con la ampliación de la oferta de aspectos de valor agregado para mejorar el sitio web de cada revista.

**PALABRAS CLAVE**

Revistas científico-académicas; calidad editorial; Acceso Abierto; comunicación científica; Universidad Nacional; Costa Rica.

## Referencias

UNIVERSIDAD NACIONAL, COSTA RICA. (2022). Portal de revistas académicas de la Universidad Nacional, Costa Rica. <https://www.revistas.una.ac.cr/>

## [Ver póster](#)

# RedELCA: una propuesta para visibilizar la producción literaria y cultural en Centroamérica y el Caribe

**NATALIA CARBALLO MURILLO**

Universidad de Costa Rica (Costa Rica)

**SUJEY SOLERA AGUILAR**

Universidad de Costa Rica (Costa Rica)

## RESUMEN

Una de las principales debilidades que presentan a nivel regional las estrategias de difusión y referencia de los estudios literarios y culturales centroamericanos, para que sean parte de una comunicación científica y cultural abierta, radica en la falta de una inter-conexión entre sus diferentes centros y programas de investigación, maestrías académicas y departamentos de enseñanza cuya producción científica es prácticamente desconocida a nivel regional; tal motivo llevó a la creación de la Red de investigación y referencia sobre estudios literarios y culturales centroamericanos (RedELCA) de acceso abierto. La RedELCA tiene como objetivo estimular el intercambio de la producción académica y consulta bibliográfica de acceso libre, relacionada con el ámbito de los estudios literarios y culturales centroamericanos y caribeños, para visibilizar la producción bibliográfica e investigación crítica e historiográfica que se realiza actualmente, en Centroamérica y el Caribe. Metodológicamente, la creación de la red consta de tres etapas: 1. Elaboración de un directorio especializado de contactos; 2. Elaboración digital de una base de datos de carácter consultivo; 3. Publicación de un boletín digital. La población

de estudio se refiere principalmente a la comunidad académica, y comunidad en general, interesada en los estudios literarios y culturales centroamericanos y caribeños.

**PALABRAS CLAVE**

Producción literaria; cultura; redes.

*Literary production; culture; networks.*

[Ver póster](#)

# Comunicación pública de la ciencia en la Universidad de Cienfuegos: políticas y acciones

**MAYRE BARCELÓ HIDALGO**

Universidad de Cienfuegos «Carlos Rafael Rodríguez» (Cuba)

**MARIANELA DÁVILA LORENZO**

Universidad de Cienfuegos «Carlos Rafael Rodríguez» (Cuba)

## RESUMEN

La universidad tiene un papel decisivo en la producción de conocimiento científico y tecnológico, por lo cual debe asumir un compromiso responsable en su divulgación y democratización. En tal sentido, este trabajo presenta un estudio sobre las políticas y acciones de comunicación pública de la ciencia, generadas en la Universidad de Cienfuegos «Carlos Rafael Rodríguez». Se expone el análisis de contenido de algunos documentos institucionales vinculados a la temática y de las publicaciones generadas en servicios de información digital y redes sociales, que permiten identificar los lineamientos relacionados con la comunicación de la ciencia. Se analizan las entrevistas semiestructurada realizadas a ocho informantes clave de la institución; para luego describir y cotejar las acciones de comunicación pública de la ciencia, desarrolladas en diferentes ámbitos. Como resultado de esta indagación previa, se identifican diversas percepciones en relación a la comunicación del conocimiento científico, dificultades en los canales de comunicación utilizados para ello y la realización de acciones básicamente de carácter transmisivo y unidireccional de información. Finalmente muestra en síntesis una

propuesta de intervención para la mejora de la visibilidad y el posicionamiento de la actividad científica e investigativa de la Universidad de Cienfuegos.

**PALABRAS CLAVE**

Comunicación pública de la ciencia; política científica; universidad.

[Ver póster](#)

# Gestión de la investigación y visibilidad de la producción científica y normativa en el Instituto de Evaluación de Tecnologías en Salud e Investigación IETSI, EsSalud. Lima-Perú, 2021-2022

## **CRISTIAN DÍAZ VÉLEZ**

Instituto de Evaluación de Tecnologías en Salud e Investigación IETSI, EsSalud (Perú)

## **MOISÉS APOLAYA SEGURA**

Instituto de Evaluación de Tecnologías en Salud e Investigación IETSI, EsSalud (Perú)

## **JORGE L. MAGUIÑA**

Instituto de Evaluación de Tecnologías en Salud e Investigación IETSI, EsSalud (Perú)

## **PERCY HERRERA AÑAZCO**

Instituto de Evaluación de Tecnologías en Salud e Investigación IETSI, EsSalud (Perú)

## **EDEN GALÁN RODAS**

Instituto de Evaluación de Tecnologías en Salud e Investigación IETSI, EsSalud (Perú)

### **RESUMEN**

La pandemia por el virus SARS-CoV-2 (COVID-19) ha representado tremendos desafíos, pero constituido a su vez una gran oportunidad para el desarrollo de la investigación en salud. Es el caso del Instituto de Evaluación de Tecnologías en Salud e Investigación (IETSI), gerencia responsable de la evaluación de las tecnologías de salud, así como de proponer las normas y estrategias para la innovación científico-

tecnológica, la promoción y regulación de la investigación en el Seguro Social de Salud (EsSalud) del Perú. La presente sistematización aborda las estrategias y acciones implementadas en los últimos quince meses para mejorar el ecosistema institucional de investigación, entre las que destacan: suscripción a herramientas digitales para la investigación y su respectiva difusión a través de cursos, conferencias, talleres y pasantías; programa de *mentoring* para acompañamiento a profesionales de la salud en el proceso de redacción y publicación en revistas indexadas, programa Two for Research para entrenamiento de estudiantes de medicina avanzados, versión *web responsive* para mayor visibilidad de las publicaciones y la normalización de la filiación institucional en las publicaciones científicas, logrando un incremento de la producción científica institucional en SCOPUS y WoS, y la consecuente ubicación en rankings internacionales.

**PALABRAS CLAVE**

Investigación; Innovación; gestión de la investigación; administración científica.

[Ver póster](#)

# Ciencia Abierta en instituciones de educación superior estatal: estrategias y experiencias de la Subcomisión de Conocimiento Abierto del Consejo Nacional de Rectores (CONARE) de Costa Rica

## **ANDREA MARÍN CAMPOS**

Universidad de Costa Rica (Costa Rica)

## **EDWARD ARAYA HIDALGO**

Consejo Nacional de Rectores (CONARE) (Costa Rica)

## **SHARLÍN SÁNCHEZ ESPINOZA**

Consejo Nacional de Rectores (CONARE) (Costa Rica)

## **ALEXA RAMÍREZ VEGA**

Instituto Tecnológico de Costa Rica (Costa Rica)

## **ANDREA MORA CAMPOS**

Universidad Nacional (Costa Rica)

## **CAROLINA SEAS**

Universidad Estatal a Distancia (Costa Rica)

## **STEVEN SEGURA JIMÉNEZ**

Universidad Estatal a Distancia (Costa Rica)

**RESUMEN**

El objetivo de este póster es sistematizar las distintas estrategias y experiencias intra e inter institucionales de promoción de la Ciencia Abierta (CA) en las instituciones de educación superior estatal, desarrolladas desde el trabajo de la Subcomisión de Conocimiento Abierto del Consejo Nacional de Rectores (CONARE) de Costa Rica.

**PALABRAS CLAVE**

Ciencia abierta; visibilidad académica; acceso abierto; conocimiento abierto.

[Ver póster](#)

# Acceso al patrimonio cultural digital el caso del Repositorio de Patrimonio Cultural Centroamericano

**CHRISTIAN DAVID ZAMORA RODRÍGUEZ**

Universidad de Costa Rica (Costa Rica)

**SOFIA VINDAS SOLANO**

Universidad de Costa Rica (Costa Rica)

## RESUMEN

Objetivo: Conocer las dinámicas en accesibilidad que ha logrado cimentar el RCPC por medio de su plataforma web. Metodología: Posicionar al repositorio como una eficiente y estimulante herramienta pedagógica así como un espacio para el goce y ocio de la cultura centroamericana. Medir el impacto en accesibilidad que generamos, por medio de la herramienta Google Analytics; así hemos logrado sistematizar los datos en acceso al acervo de patrimonio cultural que posee nuestra página web.

## PALABRAS CLAVE

Patrimonio cultural; repositorio; arte; Acceso Abierto; difusión; interoperabilidad; Centroamérica.

[Ver póster](#)

# Planificación e implementación del Repositorio de Datos Académicos de la Universidad Nacional de Rosario (RDA-UNR): una metodología colaborativa e interdisciplinaria

## **PAOLA CAROLINA BONGIOVANI**

Unidad de Gestión de Acceso Abierto, Universidad Nacional de Rosario (UNR) (Argentina)

## **ESTEVA MARIA**

Centro Avanzado de Computación (TACC), Universidad de Texas en Austin (Estados Unidos)

## **FERNANDO DÍAZ PACÍFICO**

Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario (UNR) (Argentina)

## **ANA PAULA GUTIERREZ**

Facultad de Ciencia Política y Relaciones Internacionales, Universidad Nacional de Rosario (UNR) (Argentina)

## **LAURA BALPARDA**

Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario (UNR) (Argentina)

## **DIEGO MESTRE**

Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario (UNR) (Argentina)

## **ANALÍA SALAZAR**

Unidad de Gestión de Acceso Abierto, Universidad Nacional de Rosario (UNR) (Argentina)

## PAULINA FREAN

Unidad de Gestión de Acceso Abierto, Universidad Nacional de Rosario (UNR) (Argentina)

### RESUMEN

Se presenta una metodología colaborativa interdisciplinaria para la implementación de un repositorio Universidad Nacional de Rosario. Problema: Para avanzar en el proceso de ciencia abierta, la UNR necesita contar con un repositorio de datos y servicios de curación de datos de investigación. Una encuesta preliminar a la comunidad académica indicó que el 86% nunca había publicado datos de investigación. Asimismo, no había suficientes bibliotecarios especializados en datos. Metodología: En 2021 se conformó un equipo de trabajo interdisciplinario compuesto por investigadores de diferentes disciplinas, bibliotecarios, e informáticos y se instaló un **Dataverse** de testeo. La metodología para implementar el repositorio se centró en este espacio y en las necesidades de entrenamiento y difusión en datos abiertos. En el **Dataverse** los investigadores y bibliotecarios aprenden a curar y a publicar datos de diferentes disciplinas. Asimismo, las políticas de uso y los servicios (**guías, entrenamiento, servicios de curación**) se adaptaron a la legislación de datos Argentina y contemplan los estándares internacionales y la continua necesidad de difusión y entrenamiento de esta comunidad. Resultados: El repositorio UNR-RDA, que ya cuenta con publicaciones ejemplares de los casos estudiados durante el proceso de implementación, se abrirá al resto de la comunidad en agosto de 2022.

**PALABRAS CLAVE**

Datos de investigación; repositorio de datos; colaboración interdisciplinaria.

*Research data; data repository; interdisciplinary collaboration.*

[Ver póster](#)

# Retos de la evaluación de la investigación: Caso de las universidades públicas costarricenses

## **ANDREA MARÍN CAMPOS**

Universidad de Costa Rica (Costa Rica)

## **ADRIANA CÉSPEDES VINDAS**

Universidad Estatal a Distancia (Costa Rica)

## **ANDREA BARRANTES ARRIETA**

Universidad Técnica Nacional (Costa Rica)

## **CARLOS MONTERO CORRALES**

Universidad Nacional (Costa Rica)

## **ROLVIN SALAS**

Instituto Tecnológico de Costa Rica (Costa Rica)

## **SHARLÍN SÁNCHEZ ESPINOZA**

Consejo Nacional de Rectores (CONARE) (Costa Rica)

## **EDWARD ARAYA HIDALGO**

Consejo Nacional de Rectores (CONARE) (Costa Rica)

## **MÓNICA CALDERÓN AGUILAR**

Universidad de Costa Rica (Costa Rica)

## **RESUMEN**

La evaluación de impacto, incidencia y calidad de la investigación es un reto para las universidades públicas costarricenses, que aparece de la necesidad de la trazabilidad, transparencia y la rendición de cuentas de la gestión pública. La importancia de este proceso radica en la medición del quehacer de la comunidad científica en beneficio de las sociedades desde todas las áreas del conocimiento y hacia todas las dimensiones: cultural, económica, política y productiva. Convirtiéndose esta medición en un factor fundamental para la mejora continua de procesos. En este póster se exponen los principales avances en la búsqueda por obtener una evaluación de impacto y calidad de investigación de y desde las universidades públicas, concibiendo la ciencia abierta como eje de estos procesos evaluativos.

## **PALABRAS CLAVE**

Evaluación de impacto; métricas alternativas; calidad de la investigación.

[Ver póster](#)

ISBN 978-950-34-2219-9



9 789503 422199