

Federación de repositorios universitarios: realidad y retos en Cuba

Mirelys Puerta-Díaz

Universidade Estadual Paulista Júlio Mesquita Filho – UNESP, E-mail: mireblue18@gmail.com

Yohannis Martí-Lahera

Universidad de La Habana, Cuba-UH, E-mail: yohannis@dict.uh.cu

Edgar Bisset Alvarez

Universidade Federal de Santa Catarina – UFSC, E-mail: ebicet@gmail.com

RESUMEN

En los últimos años la sociedad presencié un considerable aumento en la cantidad de publicaciones científicas. Grande parte de estas publicaciones son resultado de las investigaciones desarrolladas en las instituciones a la producción científica de otras universidades e incluso a la que es producida dentro de ella. Una apuesta para dar solución a esto se dió en la adopción e instalación de repositorios institucionales para universitarias, una de las principales dificultades a ser vencidas por estas instituciones es el acceso la difusión científica, delante de la proliferación de repositorios surge la necesidad de unir en federaciones los repositorios con el objetivo de aumentar el acceso a la información científica. A pesar de las tendencias universales, el desarrollo de repositorios institucionales en Cuba se mantienen funcionando como islas; delante de este problema, este trabajo propone como su objetivo principal: describir el estado actual de los repositorios institucionales cubanos para advertir los retos que supone la creación de una red federada en el Ministerio de Educación Superior. El trabajo se desarrolló desde una perspectiva metodológica cuali-cuantitativa, valiéndonos de la herramienta de validación desarrollada por OpenAire utilizada para evaluar interoperabilidad y los criterios de Aschenbrenner para evaluar federaciones de repositorios. Se tomaron como muestra 7 repositorios de las principales universidades del país. Esta investigación muestra que los repositorios de tesis de las instituciones cubanas son los que mayor nivel de madurez han logrado en su definición institucional, sin embargo de modo general existen una serie de elementos a nivel conceptual, semántico y arquitectural que deben ser resueltos para que los repositorios cubanos cumplan con los criterios mínimos y puedan constituirse en nodos de federaciones nacionales o extranjeras.

Palabras clave: Federación de repositorios institucionales. Acceso abierto. Interoperabilidad. Educación superior.

1 INTRODUCCIÓN

La asimilación y aplicación de los constantes desarrollos de las tecnologías de información y comunicación, han permitido que en las instituciones cubanas se trabaje de forma continua, en la búsqueda de métodos para elaborar, adquirir, conservar y transmitir su patrimonio intelectual, siempre, con el objetivo principal de brindar a su comunidad usuaria, servicios que les permitan preservar, difundir y dar acceso al conocimiento generado en ellas. Los proyectos de

repositorios universitarios cubanos aparecieron hace dos décadas, siendo una de las principales acciones desarrollada por la comunidad científica para compartir, difundir e acceder a las principales publicaciones científicas de los investigadores nacionales. Hoy, el escenario es favorecedor para el desarrollo y consolidación de repositorios institucionales universitarios dentro de la cultura científica abierta.

La integración de estos temas con el movimiento de acceso abierto demanda una mirada hacia la integración, estandarización, colaboración y el compartir. La apertura de los datos, la granularidad, versatilidad, extensibilidad y transparencia se instituyen como los elementos indispensables en la interoperabilidad de repositorios.

En directorios internacionales Cuba posee registrados 9 de repositorios en el *Registry of Open Access Repositories* (ROAR) y 10 en el *Directory of Open Access Repositories* (OpenDOAR), de ellos el 70% y 78%, respectivamente, son universitarios (ROAR, 2018; OpenDoar, 2018). Dspace es el software más empleado, solamente dos repositorios universitarios no lo usan. El uso de este software está potenciando por el proyecto *Strengthening the Role of ICT in Cuban Universities for the Development of Society* financiado por el *Flemish Interuniversity Council (Vlaamse Interuniversitaire Raad, VLIR)* de Bélgica. (RIVERO et al, 2016).

El número de instituciones universitarias cubanas que han desarrollado repositorios institucionales ha incrementado en la última década (CANO INCLÁN *et al.*, 2015; ZACCA GONZÁLEZ; MARTÍNEZ HERNÁNDEZ; DIEGO OLITE, 2012) aunque no todos se encuentran registrados en directorios ni desarrollados al mismo nivel (MARTI-LAHERA; PUERTA-DÍAZ, 2018). Sus principales propósitos han sido divulgar y almacenar la producción científica de sus instituciones no sólo a nivel internacional sino también en el territorio nacional.

Tabla 1 - Repositorios institucionales cubanos vinculados a universidades

Nombre de repositorio / Institución	Registros	Tipo de documento	Temáticas
DSPACE@UCLV / Universidad Central de las Villas	7201	Tesis, Artículos, Libros, Monografías, Presentaciones en congreso, Patentes y <i>Preprints</i>	Multidisciplinario
Scriptorium / Universidad de La Habana	2759	Tesis de Doctorado, Tesis de Maestría y Especialidad	Multidisciplinario
Alma / Pinar del Río University	2013	Tesis, Artículos, Libros, Capítulos de libros, Presentaciones en congreso	Multidisciplinario
ReDIUC / University of Camagüey	979	Tesis, Artículos de revista, Patentes, Libros, Memorias de congreso	Multidisciplinario
Repositorio de Tesis en Ciencias Biomédicas y de la Salud / Centro Nacional de Información de Ciencias Médicas (INFOMED)	558	Tesis de Doctorado	Biomedicina y Ciencias de la Salud
Repositorio / Universidad Tecnológica de La Habana José Antonio Echeverría	8228	Tesis de Doctorado, Tesis de Maestría	Ingeniería
Repositorio / Universidad de Matanzas	50	Tesis de Doctorado, Tesis de Maestría y Licenciatura (pregrado)	Multidisciplinario

Fuente: Elaboración propia (2018).

Los repositorios universitarios en Cuba se mantienen funcionando como islas. Acelerar esa evolución natural hacia la integración por medio de federaciones que menciona Shintaku *et al.* (2015) constituye imperativo para la educación superior cubana, sin embargo aún nos falta un “entorno abierto de repositorios para la curaduría, modelado semántico y el compartir información basada en archivos” (ASCHENBRENNER, 2009, p.5). Este trabajo se propone como objetivo describir el estado actual de los repositorios institucionales cubanos para advertir los retos que supone la creación de una red federada en el Ministerio de Educación Superior.

2 PROCEDIMIENTOS METODOLÓGICOS

La naturaleza de la investigación es cuali-cuantitativa de tipo descriptiva. Se estudia el funcionamiento de los repositorios institucionales cubanos por medio de la observación y el registro de datos de su realidad actual. Implica en su análisis tanto indicadores medibles como aquellos difíciles de medir cuantitativamente (HAGUETTE, 2003). Procura construir conocimientos empíricos actuales sobre el fenómeno, cuya materialización se presenta en tablas

comparativas de la muestra seleccionada y las hipótesis abordadas en la sección de resultados y discusión.

El análisis de contenido, como método de investigación permitió mediante una técnica de análisis categorial la identificación de un conjunto de indicadores de evaluación del estado de los repositorios institucionales de universidades cubanas. Además, el análisis de los indicadores posibilitó advertir los retos que suponen implementar un entorno abierto de repositorios y la red federada en el Ministerio de Educación Superior.

Para alcanzar este propósito se consideró como fuentes de información la *Base de Dados em Ciência da Informação Acervo de Publicações Brasileiras em Ciência da Informação* (BRAPCI) de la Universidad Federal de Paraná, la *Library & Information Science Abstracts (LISA)* y también fue empleado el buscador Google Scholar para la búsqueda de Información. Los trabajos académicos fueron recuperados a partir de la siguiente expresión de búsqueda en idioma inglés “*repository federation*” + “*institutional repository*”. Se consideraron como tipología documental artículos de revistas, libros, tesis, presentaciones en congresos, reportes, haciendo énfasis en aquellas que fueron sujetas al proceso de revisión por pares y publicadas entre el período 2010-2018.

La muestra de este estudio la conforman 7 repositorios institucionales de la Educación Superior cubana:

DSPACE@UCLV (Universidad Central de las Villas, MES)

Scriptorium (Universidad de La Habana, MES)

Alma (Universidad de Pinar del Río “Hermanos Saíz Montes de Oca” -UPR)

ReDIUC (Universidad de Camagüey, MES)

Repositorio de Tesis Doctorales para las Ciencias Biomédicas y de la Salud (Centro Nacional de Información de Ciencias Médicas (Infomed), Ministerio de Salud Pública-MINSAP)

Repositorio de la ISPJAE (UNiversidad Tecnológica de La Habana, MES)

Repositorio de UM (Universidad de Matanzas, MES)

La muestra es conformada por 7 repositorios institucionales de la Educación Superior, de los cuales 6 se encuentran adscritos al MES y en el caso específico del Repositorio de Tesis

Doctorales para las Ciencias Biomédicas y de la Salud pertenece al Ministerio de Salud Pública de Cuba. Su inclusión como parte de la muestra, a pesar de no estar dentro de la RED universitaria del MES, resulta fundamental ya que incluye todas las tesis defendidas en programas de postgrados de las Universidades de Ciencias de Médicas cubanas (estas se adscriben administrativamente al Ministerio de Salud Pública de Cuba y no al MES como la mayor parte de las instituciones de educación superior) y son pioneros en el desarrollo de repositorios y en la implementación de buenas prácticas de acceso abierto a la producción científica universitaria en el país.

También, este estudio considera como fuentes de información la plataforma de la Red Federada de Repositorios Institucionales de Publicaciones Científicas (LA Referencia), los directorios OpenDOAR y ROAR; los sitios web de los repositorios institucionales cubanos y los informes públicos sobre el desarrollo de los repositorios institucionales estudiados. LA Referencia, OPENDOAR y ROAR fueron empleados para el análisis de la situación de los repositorios institucionales latinoamericanos pues ofrece información actualizada a partir de las cosechas de datos realizadas a sus plataformas. Los sitios web oficiales y los documentos de trabajo de los repositorios institucionales de la Educación Superior cubana fueron empleados para completar la caracterización.

La base metodológica para la caracterización de los repositorios estudiados es RECOLECTA, la cual propone una guía para la evaluación de repositorios institucionales (AZORÍN MILLARUELO *et al.*, 2017). Esta metodología, desarrollada por la Fundación Española para la Ciencia y la Tecnología (FECYT) y la Red de Bibliotecas Universitarias (REBIUN), fue adaptada en el presente estudio para su aplicación al contexto cubano. La Guía está conformada por 8 secciones para el análisis de los repositorios, de ellos no fueron aplicados en este estudio los relativos a la seguridad, autenticidad e integridad de los datos porque no se encuentra disponible la información para su análisis. Tales razones condujeron a que no fueran aplicados todos los indicadores del resto de las secciones. RECOLECTA se apoya además en las directrices de OpenAIRE (Open Access Infrastructure for Research in Europe) para la validación de los metadatos de las publicaciones y la implementación de OAI-PMH. La herramienta de validación utilizada para evaluar interoperabilidad fue la desarrollada por OpenAire. A continuación se mencionan las secciones, indicadores y criterios de medida aplicados:

Tabla 2 - Criterios de la metodología RECOLECTA para evaluar los repositorios de la muestra

Criterios	Descripción
Visibilidad	<ol style="list-style-type: none"> 1. mide la presencia en directorios, recolectores nacionales e internacionales; 2. la existencia de un nombre propio y normalizado en directorios y recolectores; 3. existencia de una URL amigable y segura.
Políticas	<ol style="list-style-type: none"> 1. Existe una declaración sobre la misión y objetivos del repositorio; 2. Cuenta con documento de acceso público sobre el depósito en el repositorio y sobre la preservación de los contenidos; 3. Cuenta con una política pública sobre reutilización de metadatos; 4. Existe una oferta de contacto y asesoramiento visible 5. Existe una política institucional sobre acceso abierto
Aspectos legales	Garantiza que los contenidos están libres, como mínimo, de restricciones en su visualización y descarga
Metadatos descriptivos de la publicación	<ol style="list-style-type: none"> 1. Uso de modelos de metadatos estandarizados y protocolos de interoperabilidad 2. Cumplimiento de las reglas de metadatos y el uso de lenguajes controlados para la asignación de la materia
Nivel de Interoperabilidad	<ol style="list-style-type: none"> 1. Estado de la Interoperabilidad Sintáctica y Estructural 2. Integración con otros sistemas de información de la institución
Estadísticas	<ol style="list-style-type: none"> 1. Proporciona datos públicos 2. Ofrece Estadísticas públicas
Servicios y funcionalidades de valor añadido	<ol style="list-style-type: none"> 1. Posibilidad de enlazar con redes sociales 2. Integración del repositorio con gestores bibliográficos 3. Visualizar/exportar los metadatos en diferentes esquemas 4. Servicios de alerta 5. El repositorio ofrece información

Fuente: Elaboración propia (2018).

Por otro lado, el presente estudio también considera que el éxito de una Federación de Repositorios depende de criterios que están, según Aschenbrenner, *et. al.* (2010, p. 6), distribuidos en 3 niveles, nivel conceptual; nivel semántico (objeto) y nivel arquitectural (sistema) y que están estrechamente relacionados con el contexto de la aplicación y las

peculiaridades de los repositorios institucionales que se constituirán en nodos. Por ello optamos por considerar como criterios para advertir los retos que supone la creación de una red federada en la educación superior cubana, los patrones para la federación de repositorios, propuestos por Aschenbrenner, *et. al.* (2010, p. 6) y relatados a continuación:

Tabla 3 - Patrones para Federación de repositorios propuestos por Aschenbrenner

Patrones	Descripción
Escalabilidad	Mide el rendimiento de cada agente y el impacto que este causa a la eficiencia del entorno federado.
Consistencia	Mide la estabilidad de la federación a partir de la detección objetos digitales duplicados que causan problemas de consistencia en la federación.
Apertura	Analiza la capacidad de la federación de permitir que diversos agentes descentralizados interactuar ad hoc dentro de una federación
Estándares	Uso de estándares que permitan habilitar la apertura y, por lo tanto, la interoperabilidad en los agentes descentralizados, y también la flexibilidad para incorporar estándares con respecto a la sintaxis, la estructura o la semántica

Fuente: Elaboración propia (2018).

3 MARCO TEÓRICO

Los Repositorios Institucionales (RI) se mantienen como un tema relevante de alcance internacional, en especial, dentro el sector de las universidades (MARTÍ-LAHERA et al, 2016; NARAYAN; LUCA, 2017; BANGANI, 2018). Sin lugar a dudas el cúmulo considerable de estudios emanados de este contexto ha favorecido la madurez de su abordaje teórico y el avance en el desarrollo de los sistemas, todo lo cual ha sido condicionado a la par por su interconexión con otros dos importantes desarrollos, los relativos a las bibliotecas digitales y al Movimiento de Acceso Abierto (JANTZ; WILSON, 2008).

Lynch (2003, p.2) define un repositorio institucional de la siguiente manera:

es un conjunto de servicios que una universidad ofrece a los miembros de su comunidad para la gestión y difusión de materiales digitales creados por la

institución y sus miembros. Es esencialmente un compromiso de la organización con la gestión de estos materiales digitales, incluida la preservación a largo plazo, según corresponda, así como la organización, el acceso o la distribución.

Como término fue acuñado por la *Scholarly Publishing for Academic Resources Coalition* (SPARC), y el enfoque de su dimensión modélica trasciende, hoy día, la inicial perspectiva de repositorio como colecciones digitales (CROW, 2002) o “como servicio” (LYNCH, 2003). Los RI son sistemas que facilitan el acceso y preservación a largo plazo de la producción científica de una institución académica, una plataforma para la creación de revistas revisadas por pares (BANKIER; PERCIALI, 2008) o una solución para aumentar el número de citas (CHEN; HSIANG, 2009).

Las investigaciones muestran una paulatina y creciente transformación en los enfoques dados por la academia a la importancia y uso de los RI, por lo cual este debate teórico no está próximo a concluir. Definiciones más recientes de RI han insertado nuevas dimensiones, esta vez dentro del contexto de la gestión de información del ámbito académico, redefiniendo así, su relación con las nociones de producción, disseminación y uso de los recursos (BHARDWAJ, 2014). Siguiendo esta línea de mudanzas es importante mencionar a los autores Narayan e Luca (2017), los cuales sitúan a los RI no solo como vía para alcanzar mayor visibilidad y descubrimiento de las investigaciones de una universidad, sino lo abordan como criterio de *medida* de su prestigio en la actual *Era Digital*.

Esta realidad se hace evidente en la proliferación de iniciativas de ranking internacionales de universidades (*Webometrics, Academic Ranking of World Universities, THE World University Ranking*), donde la visibilidad y grado de citación de su producción científica funge como piedra angular. Ello potencia la voluntad de estas instituciones por mejorar su posición en los rankings académicos, los cuales presentan como elemento común de sus indicadores de análisis las iniciativas Open Access (en el caso específico del ranking de *Webometric*), la interoperabilidad y la promoción de las publicaciones científicas. PUYUELO (2016) afirma que el auge del Movimiento de Acceso Abierto a la Ciencia y la cada vez mayor importancia de su posicionamiento y rendimiento en espacios web universitarios, convierte a los repositorios universitarios en una herramienta clave para la comunicación y visibilidad de su producción académica y científica.

Estas condiciones contextuales han propiciado la emergencia de un nuevo paradigma en la gestión y desarrollo de los RI, considerando su gran potencial para funcionar como plataforma colaborativa, las iniciativas centralizadas se han unido en federaciones de repositorios para compartir y tener acceso mutuo a sus recursos. (LIAUWN, 2011; TABARES et al., 2015)

La necesidad de federar repositorios surge en dos escenarios distintos: (VAN DE SOMPEL; CHUTE, 2008).

Relacionado con los problemas de escalabilidad en su funcionamiento, lo cual exige la continuidad del servicio a través del paralelismo de la infraestructura de los repositorios.

La administración de los múltiples repositorios distribuidos que contienen colecciones de interés para determinadas comunidades o aplicaciones, en tanto una federación ofrece un enfoque unificado para su representación.

Las federaciones como término se han abordado desde las perspectivas disciplinares de Ciencias Políticas, Ciencia de la Computación, entre otras; en relación a los postulados de la Ciencia de la Información Weitzel (2006, p. 143) considera a las “federaciones como fuentes terciarias”. Shintaku (2014) revela las federaciones como sistemas que agregan informaciones oriundas de distintos sistemas de información. Por lo cual los repositorios federados, partiendo de esta perspectiva, tienen una relación cercana a la oferta de servicios, a las funciones de compartir y controlar los accesos a la información, aproximándose a los sistemas de información. Por su parte Tabares et al. (2015) una federación ofrece un enfoque unificado y uniforme de representación de los repositorios, tanto en visibilidad como en gestión, definiendo un sistema jerárquico que centraliza el acceso e información en un solo portal. Ambos autores afirman que la interoperabilidad es un aspecto inherente a la federación, en este sentido el Protocolo de la Iniciativa de Archivos Abiertos para la Recolección de Metadatos (OAI-PMH) provee la dimensión técnica básica para desarrollar un modelo de repositorio federado. Acorde con este planteamiento Coneglian e Santarém Segundo (2016) afirman que la interoperabilidad es una característica que permite a los usuarios acceder a objetos digitales de diversos repositorios en un único ambiente.

A partir de estos análisis se hace necesario avizorar qué capacidades básicas debe tener un modelo de federación de repositorios. Rehak *et al.* (2005) proponen los siguientes:

los contenidos publicados deben estar ampliamente disponibles,

los contenidos sean fácilmente descubiertos,
las informaciones sobre responsabilidad, derechos, accesos, procedencia y persistencia puedan ser gestionadas,
la existencia de mecanismos estándar para el acceso a los contenidos.

Lo anterior nos conduce a afirmar que sin un sistema de políticas no es posible desarrollar adecuadamente una red federada de repositorios. En una institución comprometida con el acceso abierto hará inevitable que ese sistema se componga por políticas de acceso abierto. Sin embargo, Shintaku (2014) afirma que en las iniciativas de federación apenas se consideran la publicación de datos abiertos a pesar de que estas están estrechamente vinculada al movimiento de acceso abierto.

En el nuevo sitio oficial del directorio OpenDOAR (2018) se refiere la existencia de 5 tipos fundamentales de políticas de acceso abierto que se deben declarar en el desarrollo de los repositorios. Estas son políticas de:

- a) Metadatos que define el acceso y re-uso de la descripción que se provee de los ítems del repositorio.
- b) Datos que define el acceso y re-uso a los ítems completos.
- c) Contenido que define el tipo de repositorio, los tipos de materiales que se pueden depositar y los idiomas principales.
- d) Envío concerniente a quién puede depositar, cuáles son las normas de depósito, procedimientos de moderación, control de calidad, embargos y derechos de autor.
- e) Preservación incluye los períodos de retención, la preservación funcional, preservación de ficheros, política de retiro/eliminación de ítems, control de versiones y cierre.

Precisamente en la política de preservación es donde se encuentran los principales desafíos para instituciones con bajos presupuestos e infraestructura poco desarrollado. La credibilidad y fidelización de los usuarios en los servicios que se proveen a partir de los repositorios descansan sobre la sostenibilidad y la capacidad de recuperación del sistema ante desastres.

McDonald e Walters (2011) afirman que la pérdida de datos en sistemas complejos, ya sea por desastres naturales o más probablemente por un error humano, es inevitable. De aquí se infiere que el diseño e implementación de redes federadas debe sustentarse en la teoría del

cambio y en la adecuada identificación y gestión de los riesgos con alta probabilidad de ocurrencia que enfrenta el proyecto. Por ello se instituyen como requisitos imprescindibles la inclusión de aspectos que regulen los procedimientos operativos estándar como los casos de desastres y pérdida de datos.

4 DISCUSIÓN Y RESULTADOS

En relación al indicador de visibilidad se identificó que:

Los repositorios de tesis de las universidades de Matanzas y Tecnológica de La Habana no tienen nombres propios.

El Repositorio de Tesis en Ciencias Biomédicas y de la Salud no posee un nombre abreviado ni declara oficialmente un acrónimo. Su url es amigable (tesis.sld.cu) aunque no incluye su nombre.

Se observó que no se incluye el nombre propio del Repositorio de la el Universidad de Pinar del Río en los registros de OpenDOAR, ROAR y ni en la solicitud a OpenAIRE.

El Repositorio de Tesis en Ciencias Biomédicas y de la Salud es el de mayor visibilidad. Aparece en OpenDoar, ROAR, OpenAIRE, BASE y Google académico.

Solamente Scriptorium de la Universidad de La Habana presenta una url segura mediante protocolo https.

Scriptorium y el Repositorio de Tesis en Ciencias Biomédicas y de la Salud ofrecen información sobre el depósito de documentos.

Los equipos de trabajo de Scriptorium, Dspace@UCLV y del Repositorio+ de Tesis en Ciencias Biomédicas y de la Salud han desarrollado talleres científicos para fomentar la cultura de acceso abierto en su comunidad.

Scriptorium se destaca por la utilización de las redes sociales para promover el uso del repositorio y su importancia.

Todos los repositorios institucionales ofrecen más del 75% de los textos completos en acceso abierto.

Los criterios de organización de los contenidos en los diferentes niveles de agrupación (comunidades, subcomunidades y colecciones) varían de un repositorio a otro. En la muestra estudiada se pueden encontrar desde: 1) el patrón obvio que refleja la estructura organizacional interna de la universidad; 2) el patrón de categorización a partir de los tipos de contenidos; 3) el

patrón de categorización por encabezamientos de materia; 4) el patrón de ordenación cronológica, hasta; 5) la estructuración de contenidos no consistente con ninguno de los mencionados donde en el primer nivel de agrupación se pueden encontrar varios criterios empleados sin un patrón definido. Solamente el Repositorio de Tesis en Ciencias Biomédicas y de la Salud aplica un vocabulario controlado, el Tesoro de Medline (Medical Subject Headings, MeSH).

En general, el desarrollo de las políticas y los aspectos legales de los repositorios universitarios cubanos es insuficiente. Todos declaran sus misiones y objetivos en sus sitios oficiales, sin embargo en el directorio OpenDoar ninguno ha registrado información sobre sus políticas. Al analizar los sitios oficiales se pudo constatar que en la mayoría de los repositorios cubanos permiten acceso al texto completo y descarga gratis al 75% o más de sus contenidos; sin embargo tienen indefinidos los aspectos legales y las políticas sobre re-uso de metadatos, datos sobre los textos completos y preservación. Tampoco se abordan los temas relacionados a la estandarización ni la implementación de mecanismo de validación de metadatos. El mayor desarrollo en aspectos legales y políticas lo poseen Scriptorium y el Repositorio de Tesis en Ciencias Biomédicas y de la Salud, llegando en este último a incluirse la política de preservación digital. En todos es posible encontrar información de contacto institucional de sus desarrolladores.

Las garantías y gestión de riesgos no son aspectos incluidos en la información de acceso público en los sitios oficiales ni en los documentos publicados en revistas sobre la implementación de los repositorios estudiados. Solamente un repositorio no emplea el esquema OAI-DC y en aquellos donde se emplea se completan los campos obligatorios de las directrices OpenAIRE 3.0.

La verificación del uso del OAI-PMH indican que el Repositorio de Tesis en Ciencias Biomédicas y de la Salud es único con el que se obtienen pruebas de validación exitosas en relación a la interoperabilidad. Coincide con que este es el único repositorio registrado en OpenAIRE.

El estudio de los sitios oficiales arrojó que explícitamente ningún repositorio establece la integración con otros sistemas de información de la institución. Tampoco se ofrecen estadísticas de forma pública sobre número de ítems en acceso abierto, accesos y descargas de forma

individualizada para cada ítem almacenado. Rediuc de Camagüey y Alma de Pinar del Río tienen activada la funcionalidad de estadísticas por ítem para usuario autenticados.

Las estadísticas que se ofrecen en la mayoría de los casos es sobre la evolución de los contenidos. El Repositorio de Tesis en Ciencias Biomédicas y de la Salud es el único que no las proporciona.

Se detectó, respecto a Servicios y funcionalidades de valor añadido, que los repositorios estudiados no se vinculan a redes sociales ni gestores bibliográficos. Tampoco ofrece la funcionalidad de metadatos en diferentes esquemas.

Las secciones dedicadas a la ayuda se circunscriben, en la mayoría de los casos a cuestiones técnicas de la infraestructura de la plataforma. El Repositorio de Tesis en Ciencias Biomédicas, Scriptorium y Alma además ofrecen servicios de alerta, preguntas frecuentes y enlaces externos hacia otras fuentes de información científica.

Al analizar como situación de partida el estado descrito de los repositorios se puede afirmar que la mayoría de los repositorios no están listos para incluirse en agregadores, directorios ni redes federadas. Los principales retos que enfrenta un proyecto de federación de repositorios institucionales en la educación superior cubana son:

- Si bien favorece el uso del protocolo OAI debe reconocerse que uno de los principales retos está en la identificación de duplicados. Esta duplicidad puede estar dada por varias razones entre las que fundamentalmente se encuentran la colaboración interinstitucional a nivel nacional y la participación en programas académicos de postgrado en instituciones diferentes a la de afiliación del autor de los contenidos. Hasta la actualidad, la compra y el uso de DOIs no es una posibilidad para las instituciones cubanas por el embargo económico; también atenta el bajo nivel de uso de los ORCID personales u otro identificador de investigador.
- El enriquecimiento de metadatos aumenta el valor de la federación de repositorios institucionales y llevarlo a vías de hecho implica no solo determinados retos tecnológicos, sino también en el orden político, de estrategias organizacionales y también en el plano económico para hacerlo sustentable. Si bien su implementación no conlleva reemplazar la infraestructura de gestión de datos existente, el soporte efectivo para este enriquecimiento semántico requiere de determinados procesos de representación que soporten el acceso, descubrimiento, asociación, enlazado y reusabilidad de los datos. El proceso de modelado de datos implica en un primer momento identificar los recursos importantes contenidos en los

repositorios, cumpliendo con su estructura jerárquica y propiedades. Un desafío para el profesional resulta en la búsqueda de vocabularios como nombres alternativos para describir adecuadamente los campos de recursos definidos en el esquema de metadatos empleados junto al protocolo OAI. También es muy importante diseñar URIs únicos y persistentes de forma que mantenga intacta la información de procedencia de los datos. La interconexión con conjuntos de datos enlazados en fuentes externas es un paso crucial. Más allá de los retos que implica cumplir con los principios de publicar datos abiertos enlazados y enriquecidos semánticamente, se encuentra implementar en la plataforma la búsqueda de información basada en relaciones semánticas que garanticen una óptima experiencia de usuario. Para el funcionamiento adecuado de la federación este enriquecimiento debe estar basado en el uso de vocabularios controlados que mejoren la calidad y consistencia de los metadatos de entrada.

- Otro desafío es garantizar la apertura e interoperabilidad basadas en una adecuada plataforma legal y política de cada uno de los nodos de la red. Los responsables de implementar repositorios deben estar conscientes de la centralidad de los temas vinculados al copyright y a los derechos de propiedad intelectual. Uno de los mayores problemas a resolver por los gestores de repositorios estudiados es la posición indefinida que mantienen dentro del ecosistema de publicación. Esa postura no debe quedar clara sólo para el autor de los contenidos en el acuerdo que tiene con la institución huésped. También debe quedar claramente explicitada para otros editores, de otra manera se están limitando las posibilidades de publicación de obras derivadas (por ejemplo de tesis o documentos de trabajo) en canales formales de comunicación científica que exigen obras inéditas. Además de lo anterior, los gestores de repositorios deben prever la asesoría y adecuada información a los autores sobre la inclusión de trabajos publicados con anterioridad en otras fuentes para evitar conflictos legales. Los autores y gestores de repositorios deben estar familiarizados con las políticas de autoarchivo y cuáles son las que se aplican en las fuentes primarias de publicación de los trabajos. Otra de las aristas a considerar para mitigar los riesgos, consolidar la confianza de los miembros de la red y asegurar los derechos de autor es explicitar a los usuarios finales del sistema qué pueden o no hacer con los contenidos y metadatos que se encuentran a su disposición.

- La preservación digital constituye uno de elementos más sensibles en la sostenibilidad del proyecto de federación de repositorios. La adaptación a tecnologías cambiantes, la adopción de tecnologías emergentes, disruptivas resulta complejo para países en vías de desarrollo

como Cuba. Esta situación complejiza por tanto la preservación a largo plazo de contenidos digitales. Otra cuestión asociada a este tema y en estrecho vínculo con los aspectos legales es la necesidad de insertar en los acuerdos que se suscriben con el/los propietario (s) de los contenidos la posibilidad de modificar formatos con el propósito expreso de la preservación a largo plazo y la estandarización, más allá de la obsolescencia de los soportes o los softwares que aseguran la disponibilidad y accesibilidad a los contenidos. La preservación a largo plazo exige la acumulación y perdurabilidad que también implican que la infraestructura de la red federada debe ser escalable. El equipo de desarrollo de la red debe ser capaz de modelar de manera confiable las tasas de aumento de datos y gestionar la tecnología necesaria para escalar los requisitos de almacenamiento necesarios.

- El control de la calidad, en estrecho vínculo con la preservación, resulta crucial para asegurar tres aspectos claves: 1) la integridad académica, 2) la integridad de los ficheros y 3) el cumplimiento de los estándares. Constituyen desafíos el diseño e implementación de un modelo que asegure la ejecución rigurosa de tareas rutinarias, controles internos y auditorías. El cumplimiento de estándares no sólo facilita la indización de repositorios en otras bases de datos y directorios de prestigio internacional sino que también aumenta la facilidad de integración en el contexto informacional. La calidad de los contenidos y de los metadatos de la red federada hará que sea más fácil la articulación tanto desde el punto de vista tecnológico como cultural. Otro elemento relevante en este desafío resultan las tareas de la curaduría de contenidos que tributan a la consistencia y estandarización dentro de la red. Las actividades de curación para la federación en Cuba se tornarán desafiantes en tanto se ha podido constatar que casi todos los repositorios estudiados no se trabaja con vocabularios controlados (en sus más amplia acepción), ni emplean los mismos criterios de organización de sus contenidos, tampoco poseen políticas definidas de datos, metadatos, preservación. Esto sugiere la necesidad de facilitar herramientas que antes de la cosecha validen los envíos. Esta validación deberá centrarse en garantizar que todos los datos vayan acompañados de suficientes metadatos administrativos, descriptivos, estructurales y técnicos y el ajuste a las políticas de la federación.

- La creación y fortalecimiento de capacidades es esencial. Se instituye como factor crítico de éxito un grupo multidisciplinario cuyos objetivos de participación, roles en la cosecha, evaluación, enriquecimiento, uso y disseminación queden claramente definidos. Este grupo debe participar activamente en la búsqueda de soluciones y la creación de las capacidades correspondientes para incorporar a las instituciones de educación superior que no han logrado

establecer la infraestructura básica ni el contexto legal y/o político de su repositorio institucional. El reto de formar una comunidad de práctica representativa de varias disciplinas debe ser satisfactoriamente superado para lograr que las políticas, reglas, plataformas, elecciones tecnológicas sean apropiadas a la realidad nacional y ajustadas a los estándares internacionales para evitar que la federación sea una isla difícil de interoperar con otras iniciativas foráneas.

5 CONSIDERACIONES FINALES

El desarrollo de una iniciativa de repositorios federados propicia la emergencia de un nuevo modelo de acceso y uso a la información académica y científica de una sociedad. La federación de repositorios institucionales en el Ministerio de Educación Superior cubano, apoyada en el uso de las tendencias infotecnológicas actuales (sistemas de recomendación, computación en la nube, arquitectura responsive, sincronización con las redes sociales), diversificará sus ofertas informacionales y el enriquecimiento de la experiencia de los usuarios, permitiendo servicios de valor agregado como la vinculación de citas, la creación de interfaces de consulta entre los repositorios miembros, así como brindar servicios de recomendación federado.

La caracterización de los repositorios estudiados muestra que los repositorios de tesis son los que mayor nivel de madurez han logrado en sus definición institucional, apertura e interoperabilidad y en la definición de políticas y aspectos legales.

Una vez que existen en el país varios repositorios institucionalizados, el próximo paso es crear una red federada a nivel nacional que allane el camino para participar como nación en otras federaciones regionales. La federación nacional fortalece las capacidades nacionales de enseñanza-aprendizaje e investigación. Un grupo de repositorios institucionales interconectados y dispersos en la geografía nacional traerán como resultado el aumento de la disponibilidad y accesibilidad de recursos para la enseñanza y la investigación. También fomenta la colaboración nacional, la interdisciplinariedad y el mapeo de expertos.

La creación de una red federada de repositorios universitarios cubanos supone desafíos relacionados con la gestión de información. Estos desafíos se evidencian en los 3 niveles: conceptual, semántico (objeto) y arquitectural (sistema); así como en los patrones de escalabilidad, consistencia, apertura y estandarización. Lo que permitiría pensar en una red que a nivel semántico y técnico, sea lo suficientemente flexible como para capturar metadatos que le permita manejar una amplia variedad de recursos con disímiles niveles de acceso y términos de uso.

Las federaciones de repositorios provee un punto de acceso y búsqueda de sus recursos informacionales disponibilizados de forma centralizada y enriquecida en una única plataforma, propiciando el descubrimiento de la información en beneficio de la experiencia de sus usuarios, todo ello en base a la cosecha de metadatos de los repositorios contribuyentes. Estas características generan el imperativo del desarrollo de un sistema cuya arquitectura cumpla con los requisitos técnicos (escalabilidad, apertura, estandarización y consistencia) para hacer de la federación una realidad sostenible.

Contar con una red federada de repositorios institucionales de las universidades cubanas, posibilitará el control sobre los trabajos de investigación, los objetos de aprendizajes, las tesis, los conjuntos de datos y otros documentos generados por la academia Cuba. Siendo así, la postura indefinida que mantienen dentro del ecosistema de publicación los repositorios estudiados no debe ser heredada por la red federada si estos se constituyen nodos de ella. El establecimiento de la red federada de repositorios debe partir de una sólida y bien conformada plataforma política y legal. Estos elementos son claves para la sostenibilidad del proyecto.

Un proyecto de esta naturaleza y propósitos tributa a la mitigación de disparidades en las capacidades tecnológicas y de recursos humanos en las instituciones de educación superior. La conformación de una comunidad de práctica a nivel nacional y las alianzas son vías expeditas para lograr la incorporación de aquellas instituciones, que a pesar de tener voluntad política no poseen los recursos suficientes para incorporarse a la red. Estas opciones harán la red más inclusiva y representativa de la producción científica de la academia cubana.

Aunque que la red federada de repositorios universitarios no es “la solución” definitiva ni completa que necesita la educación superior cubana para la gestión integral de sus recursos digitales de información, es válido notar que sí constituye una pieza clave. La implementación de esa red, - aunada a disposiciones de autoridades universitarias de preservar el valioso patrimonio que constituye la producción intelectual de la educación superior cubana, contribuirá a reducir las brechas de visibilidad digital entre sus miembros.

Federação de repositórios universitários: realidade e desafios para Cuba

RESUMO

Nos últimos anos, a sociedade testemunhou um aumento considerável no número de publicações científicas. Grande parte dessas publicações é resultado de pesquisas realizadas em instituições universitárias, dentre as principais dificuldades a serem superadas por estas instituições está, o acesso à

produção científica de outras universidades, e até mesmo, àquelas que são produzidas dentro delas. A principal proposta para dar solução a este fenômeno, radicou na adoção e instalação de repositórios institucionais para o armazenamento e divulgação científica. Diante da proliferação de repositórios institucionais nas instituições de ensino superior, emerge a necessidade de criar Federações repositórios com o objetivo de aumentar o acesso à informação científica. Apesar das tendências universais, o desenvolvimento de repositórios institucionais em Cuba continua a funcionar como ilhas; diante deste problema, este trabalho propõe como objetivo principal: descrever o estado atual dos repositórios institucionais cubanos para alertar sobre os desafios que supõe a criação de uma rede federada no Ministério da Educação Superior. O trabalho foi desenvolvido a partir de uma perspectiva metodológica qualitativa-quantitativa, utilizando a ferramenta de validação desenvolvida com base no *OpenAire*, RECOLETA, que propõe uma guia para avaliar repositórios institucionais, assim como, alguns dos critérios apresentado por Aschenbrenner para avaliar as federações de repositórios. Tomou-se como amostra os 7 repositórios das principais universidades do país. Esta pesquisa mostra que os repositórios de tese das instituições cubanas são os que têm alcançado maior maturidade em sua definição institucional. No entanto, há uma série de elementos conceituais, semânticos e de nível arquitectural que devem ser resolvidos para que, os repositórios cubanos cumpram os critérios mínimos para formarem parte nós de federações nacionais ou estrangeiras.

Palavras-chave: Federação de repositórios institucionais. Acesso aberto. Interoperabilidade. Ensino superior.

Federation of university repositories: reality and challenges in Cuba

ABSTRACT

In recent years, society has witnessed a considerable increase in the scientific publication's number. Most of these publications are the result of research carried out in universities. Among the main difficulties to be overcome by these institutions is access to scientific literature from other universities and even those that are produced within them. A bet to solve this, it occurred in the adoption and installation of institutional repositories for scientific diffusion, opposite the proliferation of institutional repositories, there is a need to join federations' repositories with the objective of increasing access to scientific information. Despite the universal trends, the development of institutional repositories in Cuba continues to function as islands. Faced with this problem, this paper proposes the main objective: to describe the current state of Cuban institutional repositories to warn of the challenges that suppose the create a repository federated network in the Ministry of Higher Education. The work was developed from a qualitative-quantitative methodological perspective, using the validation tool developed on the basis of OpenAire, RECOLETA, which proposes a guide to evaluate institutional repositories, as well, as some of the criteria presented by Aschenbrenner to evaluate the federations of repositories. The 7 repositories of the main universities of the country. This research shows how the theses repositories of the Cuban institutions present a higher level of maturity in their institutional definition, however, there are usually a series of conceptual, semantic and architectural elements that must be solved, so that Cuban repositories comply with the minimum criteria, for form part of national or foreign repositories federations.

Keywords: Federation of Institutional repositories. Open access. Interoperability. Higher Education.

REFERENCIAS

- ASCHENBRENNER, Andreas et al. Towards an Open Repository Environment. **Journal of Digital Information**, v. 11, n. 1, 2010. Disponível em: <https://pdfs.semanticscholar.org/9377/d3aa61d8ad93c927166ce5d7807ae4a5c733.pdf>. Acesso em: 19 jul. 2018.
- BANGANI, S. The impact of electronic theses and dissertations: a study of the institutional repository of a university in South Africa. **Scientometrics**, v. 115, n. 1, p. 131-151, 2018. Disponível em: <https://link.springer.com/article/10.1007/s11192-018-2657-2>. Acesso em: 02 ago. 2018.
- BHARDWAJ, Raj Kumar. Institutional repository literature: A bibliometric analysis. **Science & Technology Libraries**, v. 33, n. 2, p. 185-202, 2014. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/0194262X.2014.906018?journalCode=wstl20>. Acesso em: 02 ago. 2018.
- CANO INCLÁN, Anisley, Raiza Ana de Dios Arias, Ognara García García & Floriselda Cuesta Rodríguez. Los repositorios institucionales: situación actual a nivel internacional, latinoamericano y en Cuba. **Revista Cubana de Información en Ciencias de la Salud**, v. 26, n.4, 2015. Disponível em: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2307-21132015000400002. Acesso em: 02 ago. 2018.
- CHEN, Kuang-hua; HSIANG, Jieh. The unique approach to institutional repository: Practice of National Taiwan University. **The electronic library**, v. 27, n. 2, p. 204-221, p. 2009. Disponível em: <https://www.emeraldinsight.com/doi/abs/10.1108/02640470910947566>. Acesso em: 02 ago. 2018.
- CONEGLIAN, C. S.; SEGUNDO, J. E. S. Europeana no linked open data: conceitos de web semântica na dimensão aplicada das humanidades digitais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 22, n. 48, 2017. Disponível em: <http://www.brapci.inf.br/v/a/22274>. Acesso em: 19 Set. 2018.
- CROW, Raym, et al. **The case for institutional repositories: a SPARC position paper**. 2002.
- HAGUETTE, Teresa Maria Frota. **Metodologias qualitativas na sociologia**. 10.ed. Petropolis: Vozes, 2003. 224 p.
- KOLER-POVH, Teja; MIKOŠ, Matjaž; TURK, Goran. Institutional repository as an important part of scholarly communication. **Library Hi Tech**, v. 32, n. 3, p. 423-434, 2014. Disponível em: <https://www.emeraldinsight.com/doi/abs/10.1108/LHT-10-2013-0146>. Acesso em: 02 ago. 2018.
- LAGOZE, Carl; VAN DE SOMPEL, Herbert. The Open Archives Initiative: Building a low-barrier interoperability framework. En **Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries**. ACM, p. 54-62, 2001.
- LEWIS, Stuart et al. Repository as a Service (RaaS). **Journal of Digital Information**, v. 13, n. 1, 2012. Disponível em: <https://researchspace.auckland.ac.nz/handle/2292/17764>. Acesso em: 19 jul. 2018.

MARTI-LAHERA, Y; PUERTA-DIAZ, M. Institutional Repositories Located at Universities: An Analysis of the Cuban Experience. **Library Trends**, Special issue dedicated to Cuba, 2018. In press.

MCDONALD, Robert H.; WALTERS, Tyler O. Restoring Trust Relationships within the Framework of Collaborative Digital Preservation Federations. **Journal of Digital Information**, v. 11, n. 1, 2010. Disponível em: <https://journals.tdl.org/jodi/index.php/jodi/article/view/757>. Acesso em: 02 ago. 2018.

NARAYAN, B; LUCA, E. Issues and challenges in researchers' adoption of open access and institutional repositories: a contextual study of a university repository. **Information Research**, v. 22, n.4, p. 1-14, 2017. Disponível em: <https://opus.lib.uts.edu.au/handle/10453/121438>. Acesso em: 02 ago. 2018.

OPENAIRE. 2015. "OpenAIRE Guidelines for Literature Repositories". Disponível em: <https://guidelines.openaire.eu/en/latest/literature/index.html>. Acesso em: 2 mar. 2018.

OPENAIRE. 2018. Disponível em: <https://www.openaire.eu/validator/welcome>. Acesso em: 20 mar. 2018

OPENDOAR. Policies Tool. 2018. Disponível em: <http://v2.sherpa.ac.uk/opendoar/policytool/> Acesso em: 20 mar. 2018

PUYUELO, Lucas Torregrosa. Análisis de la visibilidad científica con la implantación de aplicaciones de publicación en abierto. 2016.

RIVERO, Manuel Osvaldo Machado, et al. **Dspace@ UCLV: Repositorio Digital en la Universidad Central "Marta Abreu" de Las Villas**. Conference: XIV Congreso Internacional de Información, Info'2016, La Habana, 2016.

ROAR. Disponível em: <http://roar.eprints.org/>. Acesso em: 2 maio 2018

SHINTAKU, Milton; DUQUE, Claudio Gottschalg; SUAIDEN, Emir José. Federações de repositórios: conceitos, políticas, características e tendências. **Perspectivas em Ciência da Informação**, v. 20, n. 3, p. 51-66, 2015. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2358>. Acesso em: 02 ago. 2017.

SHINTAKU, Milton. **Federação de repositórios científicos: identificação, análise e proposta de modelo baseado nas tendências tecnológicas e da ciência**. 2014. xviii, 268 f., il. Tese (Doutorado em Ciência da Informação)—Universidade de Brasília, Brasília, 2014.

VAN WYK, B.; MOSTERT, J. Toward Enhanced Access to Africa's Research and Local Content: A Case Study of the Institutional Depository Project, University of Zululand, South Africa. **African Journal of Library, Archives & Information Science**, v. 21, n. 2, p. 139-151, 2011.