



IV Conferencia Internacional Biredial - ISTE
**Acesso Aberto, Preservação Digital,
Interoperabilidade, Visibilidade
e Dados Científicos**

Universidade Federal do Rio Grande do Sul
15 - 17 de octubre de 2014

IV Conferência Internacional Biredial-ISTEC

IV Conferência Internacional sobre Bibliotecas e Repositórios Digitais
(BIREDIAL 2014)

IX Simpósio Internacional de Bibliotecas Digitais (SIBD 2014)

ANAIS DAS SESSÕES TEMÁTICAS E PÔSTERS

Porto Alegre
Universidade Federal do Rio Grande do Sul

2014



IV Conferencia Internacional Biredial - ISTE
**Acesso Aberto, Preservação Digital,
Interoperabilidade, Visibilidade
e Dados Científicos**

Universidade Federal do Rio Grande do Sul
15 - 17 de octubre de 2014

ANAIS DAS SESSÕES TEMÁTICAS E PÔSTERS



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Carlos Alexandre Netto

Reitor

Rui Vicente Oppermann

Vice-Reitor

Viviane Carrion Castanho

Diretora da Biblioteca Central

Jussara Issa Musse

Diretora do Centro de Processamento de Dados

Dados Internacionais de Catalogação-na-Publicação

Conferência Internacional Biredial-ISTEC (4. : 2014 : Porto Alegre, RS).
Anais das sessões temáticas e pôsters. Porto Alegre: Universidade Federal do
Rio Grande do Sul, 2014.

ISBN 978-85-66106-41-1

1. Ciência da Informação. 2. Biblioteconomia. 3. Biblioteca universitária. I.
Universidade Federal do Rio Grande do Sul. II. *Conferência Internacional sobre
Bibliotecas e Repositórios Digitais (4. : Porto Alegre : 2014)*. III. *Simpósio
Internacional de Bibliotecas Digitais (9. : Porto alegre, 2014)*. IV. Título.



Organização e realização

Presidente

Caterina Groposo Pavão - Universidade Federal do Rio Grande do Sul, Brasil (UFRGS), Brasil

Comissão Executiva

Viviane Carrion Castanho - Universidade Federal do Rio Grande do Sul, Brasil (UFRGS), Brasil

Lorete Mattos - UFRGS

Comissão Técnico-Científica

Caterina Groposo Pavão - Coordenadora – Universidade Federal do Rio Grande do Sul (UFRGS), Brasil

Marisa Raquel De Giusti - ISTE. Universidad Nacional de La Plata, Argentina

Malgorzata Lisowska - Universidad del Rosario, Colombia

Meilyn Garro - Universidad de Costa Rica, Costa Rica

Marlem Uribe Marengo - Universidad del Norte, Colômbia

Pareceristas

Ana Maria Mielniczuk de Moura - UFRGS

Laura Rezende – UnB

Angelica Conceição Dias Miranda - FURG

Luana Sales - CNEN-IEN

Ariel Lira – UNLP

Malgorzata Lisowska - UROSARIO

Bianca Amaro – IBICT

Marisa De Giusti – UNLP

Caterina Groposo Pavão - UFRGS

Marlem Uribe Marengo - UNINORTE

Fabiano Couto Corrêa da Silva - FURG

Miguel Angel Márdero Arellano - IBICT

Fernando César Lima Leite - UnB

Rafael Port da Rocha - UFRGS

Franco Terruzzi – UNLP

Remedios Melero – CSIC

Gonzalo Villarreal – UNLP

Samile Andrea de Souza Vanz – UFRGS

Helio Kuramoto UFMG

Saray Córdoba González - UCR

Jose Texier – UNET

Sônia Elisa Caregnato – UFRGS

Humbert Garrido - UROSARIO

Zita Prates Oliveira - UFRGS



IV Conferencia Internacional Biredial - ISTE
**Acesso Aberto, Preservação Digital,
Interoperabilidade, Visibilidade
e Dados Científicos**

Universidade Federal do Rio Grande do Sul
15 - 17 de octubre de 2014

Organizan:



Patrocinan:

OURO

PRATA



Colaboran:





SUMÁRIO

Apresentação.....	7
Interoperabilidade.....	8
INDIXE: el agregador nacional de la Red Mexicana de Repositorios Institucionales-REMEDI para la interoperabilidad con redes federadas de repositorios institucionales.....	9
Configuração e desenvolvimento do acervo de recursos educacionais em saúde da UNA-SUS.....	24
A migração da biblioteca digital de teses e dissertações da Universidade Federal De Goiás (BDTD/UFG): tópicos a serem considerados.....	38
Interfaces móviles en DSPACE.....	49
Repositorios DSPACE con múltiples contextos OAI-PMH.....	58
La incorporación del registro de obras y autores de una institución: estudio de caso SEDICI-UNLP-SCOPUS.....	81
Interoperabilidad entre sistemas del Instituto Español de Oceanografía.....	94
Repositório de dados científicos.....	103
Necesidades de gestión de datos científicos en argentina. El caso de la Universidad Nacional de Rosario.....	104
Visibilidade de documentos em acesso aberto: necessidade de educação do usuário.....	122
Letras espanhol no repositório institucional de conteúdo digital UFSC.....	141
Sistema de indicadores ambientais: arcabouço para projetos de pesquisa.....	153
CHAIN-REDS DART Challenge.....	166
Preservação digital.....	184
Desafíos de la preservación digital de los archivos sonoros.....	185
Evaluación automática de preservación mediante la utilización de tareas de curación en DSpace.....	201
Produção científica sobre preservação digital: levantamento bibliográfico e documental em bases de dados.....	215
Curadoria digital de dados de pesquisa: o cenário das fundações de amparo à pesquisa no Brasil.....	244



Acesso Aberto.....	255
Acesso aberto no contexto dos repositórios digitais educacionais: o caso da rede UNA-SUS.....	256
As revistas acesso aberto e as políticas de direitos autorais: o caso de revistas na plataforma Scielo-Brasil.....	272
A preservação do patrimônio arquivístico fotográfico e o acesso aberto utilizando o <i>software</i> de descrição ICA-AtoM.....	285
Portal de periódicos da UERJ na perspectiva do acesso livre: uma proposta.....	296
Visibilidade.....	309
Os investigadores e a sua relação com o acesso aberto à produção científica: o caso do ISCTE-IUL.....	310
Servicios de apoyo a la gestión de publicaciones de acceso abierto: el caso de la Universidad De Costa Rica.....	328
Difusión de los resultados de investigación en acceso abierto de la “Biblioteca Digital de Estudios Feministas Y De Género”: Proyecto temático del sistema Nacional de bibliotecas de la Universidad Nacional de Colombia.....	342
Sessão de Pôsters.....	348
Índice de autores.....	401



Apresentação

A “**Conferência Internacional Acesso Aberto, Preservação Digital, Interoperabilidade, Visibilidade e Dados Científicos**” se realizará de 15 a 17 de outubro de 2014, no Hotel Embaixador, na cidade de Porto Alegre, Rio Grande do Sul, Brasil organizada pela Universidade Federal do Rio Grande do Sul - UFRGS com a colaboração do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e do LibLink do consórcio Ibero-americano para a Educação em Ciência e tecnologia (ISTEC), Este, que é o quarto encontro de duas conferências internacionais: *IV Conferência Internacional sobre Bibliotecas e Repositórios Digitais (BIREDIAL 2014)* e *IX Simpósio Internacional de Bibliotecas Digitais (SIBD 2014)*, se propõe a ser um espaço de apresentação e compartilhamento de experiências referentes a acesso aberto, preservação de documentos digitais em repositórios institucionais e de dados científicos.

Temas da Conferência:

Acesso Aberto: Legislação, impacto e visibilidade, modelos sustentáveis para publicações, políticas, acesso aberto e educação superior, métricas de impacto, novas tecnologias e ferramentas.

Preservação: Políticas e planos de preservação, ferramentas para preservação estratégias e implementação, curadoria de conteúdos, futuro e tendências da preservação.

Repositórios de dados primários/Data sharing: Grandes volumes de dados (Big Data), formatos de dados e metadados, reprodutibilidade e infraestrutura, aplicações para conjuntos de dados em acesso aberto.

Tecnologia e Metodologia: Interoperabilidade, metadados, avanços tecnológicos, redes.

Interoperabilidade: Interoperabilidade, políticas e padrões de metadatos, redes, novas tecnologias e ferramentas, qualidade e normalização de metadados.

Visibilidade: métricas de impacto e visibilidade dos documentos em acesso aberto, Divulgação, sensibilização dos usuários, capacitação.

A Conferência foi dirigida a Bibliotecários, Analistas de sistemas, Engenheiros, Designers, Estáticos e Comunicadores; Comunidades científicas, Observatórios, Open Data Centers e outros; Vice-reitores acadêmicos; Pesquisadores e editores de publicações científicas; Estudantes de diversas especialidades; Instituições de educação superior e centros de pesquisa.



Interoperabilidade

Interoperabilidade, políticas e padrões de metadados, redes, novas tecnologias e ferramentas,
qualidade e padronização de metadados



INDIXE: EL AGREGADOR NACIONAL DE LA RED MEXICANA DE REPOSITORIOS INSTITUCIONALES-REMERI PARA LA INTEROPERABILIDAD CON REDES FEDERADAS DE REPOSITORIOS INSTITUCIONALES

Rosalina Vázquez Tapia¹, Antonio Felipe Razo Rodríguez²

¹ Maestra en Tecnología Educativa; Directora de la Biblioteca Virtual Universitaria, Universidad Autónoma de San Luis Potosí, Niño Artillero S/N Zona Universitaria, San Luis Potosí, S.L.P., México. alinavn@uaslp.mx

² Maestro en Diseño de la Información; Profesor del Departamento de Arte, Diseño y Arquitectura. Universidad Iberoamericana Puebla. Blvd. del Niño Poblano No. 2901 Unidad Territorial Atlixcáyotl, Puebla, Pue, México. antrazo@gmail.com

RESUMEN

La Red Mexicana de Repositorios Institucionales (REMERI), tiene como objetivo crear una red interconectada de repositorios digitales de instituciones de educación superior (IES) en México, para integrar, difundir, preservar y dar visibilidad a su producción científica, académica y documental; así como también, incorporarse a redes o directorios de repositorios internacionales para fomentar la colaboración y apoyar el acceso y la divulgación de contenidos de acceso abierto. El proyecto REMERI (www.remeri.org.mx) inicia en el 2011 a partir de una iniciativa presentada por la Biblioteca Virtual de la Universidad Autónoma de San Luis Potosí, para desarrollar una Red de Repositorios Institucionales de las instituciones miembros de la Red Abierta de Bibliotecas Digitales (RABID), perteneciente a la Corporación Universitaria para el Desarrollo de Internet 2 (CUDI). Se integra un grupo de trabajo de seis instituciones mexicanas de educación superior miembros de RABID, mismas que se constituirían posteriormente como fundadoras de REMERI. Ante la necesidad de contar con una plataforma tecnológica que permitiera la creación de un nodo mexicano interoperable con la Red Federada de Repositorios Institucionales de Publicaciones Científicas LA-Referencia, los representantes por parte de México en dicho proyecto, adoptan la iniciativa de REMERI para construir un primer prototipo con base en estándares internacionales predefinidos. Para la interconexión de los repositorios institucionales y su interoperabilidad con LA-Referencia, fue necesario considerar el uso de estándares de interoperabilidad basados en DRIVER 2.0 y la implementación de servidores de metadatos con el protocolo OAI-PMH. En un primer análisis de los repositorios disponibles en México para la construcción del prototipo, se encontraron servidores no estandarizados, metadatos inconsistentes, faltantes o no estructurados y una implementación parcial o con modificaciones del protocolo OAI-PMH; además, de una diversidad de recursos, formatos y áreas de conocimiento. Adicionalmente, se establecieron los siguientes requerimientos para el desarrollo de la plataforma tecnológica de REMERI: Integración en un mismo sistema de diversas colecciones de Repositorios Institucionales; recuperación de información por medio de una interfaz de búsqueda y consulta web; indexación de la colección para mostrar resultados por relevancia; interoperabilidad de la colección



utilizando el protocolo OAI-PMH; uso y normalización de los metadatos a estándares y especificaciones Dublin Core, LA Referencia y DRIVER. El resultado fue el desarrollo de un *cosechador agregador* y sistema de consulta denominado INDIXE, el cual tiene las siguientes características: desarrollo propio basado en el uso de software *open source*, configurable y flexible; desarrollo funcional con tolerancia a errores y fácil de administrar; diseño de una interfaz accesible vía web multiplataforma y navegable para la incorporación, mantenimiento, consulta e interoperabilidad de repositorios institucionales; acceso a las colecciones de REMERI por medio de mecanismos avanzados de recopilación, búsqueda y visualización de grandes colecciones; implementación de estándares web para garantizar la usabilidad en plataformas fijas y móviles, implementación de estándares de metadatos y normalización de información para facilitar su interoperabilidad. En esta propuesta se describen los componentes y arquitectura de la plataforma INDIXE, sus servicios y posibilidades de crecimiento; así como el trabajo efectuado para lograr la interoperabilidad con LA-Referencia.

Palabras clave: Interoperabilidad, Acceso Abierto, Repositorios Institucionales.

INTRODUCCIÓN

El creciente desarrollo y la rápida diseminación de la información y del conocimiento, junto con una serie de fenómenos globales de carácter económico, social y tecnológico, propiciaron hace más de una década, el surgimiento de un nuevo paradigma de la comunicación científica denominado *Acceso Abierto* (AA), el cual promueve a partir de una serie de declaraciones y definiciones, que el conocimiento científico debe ser de acceso libre y gratuito a todo mundo, sin barreras económicas, políticas, legales o de cualquier otra índole.

El movimiento de Acceso Abierto

El concepto de AA para datos científicos, se utilizó por primera vez para compartir la colección de datos referentes al año Internacional de Geofísica (International Council of Science, 1957-1958). En el 2002, la Iniciativa de Acceso Abierto de Budapest [Budapest Open Access Initiative, (BOAI)] lanzó una campaña mundial para el AA, también llamado *Open Access*, (OA) para las publicaciones científicas. La iniciativa BOAI fue la primera en utilizar el concepto de AA y la primera en proponer las dos estrategias del AA: las revistas (llamada vía dorada) y los repositorios (llamada vía verde). En Junio de 2003, en la Declaración de Bethesda (Bethesda Statement on Open Access Publishing), surgieron la primera definición de las publicaciones de AA y el concepto de depósito o autoarchivo en repositorios. En octubre del mismo año, en la Declaración de Berlín sobre Acceso Abierto al Conocimiento en Ciencias y Humanidades (Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities) se definieron las contribuciones del AA y las dos condiciones que éstas deben cumplir.

Entre las principales definiciones del AA destacan las tres siguientes:

1. “Por *acceso abierto* a la literatura científica erudita, entendemos su disponibilidad gratuita en Internet, para que cualquier usuario la pueda leer, descargar, copiar, distribuir o imprimir, con la posibilidad de buscar o enlazar al texto completo del artículo, recorrerlo para una indexación exhaustiva, usarlo como datos para software, o utilizarlo para cualquier otro propósito legal, sin otras barreras financieras, legales o técnicas distintas de la fundamental de acceder a la propia Internet. El único límite a la reproducción y distribución de los artículos publicados, y la única función del copyright en este marco, no puede ser otra que garantizar a los autores el control sobre la integridad de su trabajo y el derecho a ser acreditados y citados” (Budapest Open Access Initiative, 2003, p. 1).
2. En la declaración Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003) se establece la siguiente definición de AA: Definimos el Acceso Abierto como una amplia fuente de conocimiento humano y patrimonio cultural aprobada por la comunidad científica. Las contribuciones del acceso abierto incluyen los resultados de la investigación científica original datos primarios y metadatos, materiales fuentes, representaciones digitales de materiales gráficos y pictóricos, y materiales eruditos en multimedia.
3. La literatura de AA es digital, en línea, libre de cargos y libre de la mayoría de las restricciones de *copyright* y licencias (Suber, 2012).

De acuerdo a Suber (2012), hay muchas formas de hacer posible el AA: páginas web personales, *blogs*, *wikis*, foros de discusión, libros electrónicos, bases de datos, *webcasts*, recursos multimedia, sindicación de contenidos RSS, páginas de noticias, entre otros. Sin embargo, los dos vehículos o rutas de AA más implementados son los repositorios y las revistas.

El movimiento Open Access (OA), utiliza el término Gold OA (AA Dorado) para la entrega a través de las Revistas; y Green OA (AA Verde) para la entrega a través de Repositorios. El concepto Self-Archiving (Autoarchivo) es la práctica de depositar por uno mismo un trabajo en un repositorio OA. Los tres términos fueron acuñados por Steven Harnard, (Suber, 2012).

Definiciones de Repositorio Institucional

El archivo de contenidos científicos en repositorios digitales constituye la vía verde del AA. En la década de los 90's se usaban los archivos de eprints (*eprint archives*) y archivos abiertos (*open archives*) para hacer referencia al archivo electrónico de contenidos científicos. Posteriormente con el movimiento de AA, se utilizó el término repositorio como sinónimo de archivo, el cual se consolidó y desplazó a los conceptos anteriores. Un repositorio digital puede definirse como un sistema en red que proporciona servicios Web sobre una colección de objetos digitales, basado en una arquitectura abierta y en el uso de estándares (Abadal, 2012).



De acuerdo a su propósito, los repositorios digitales se clasifican en dos tipos: los institucionales y los temáticos. Los Repositorios Institucionales (RI) son aquellos que almacenan, preservan y proveen acceso a la producción científica y académica de una institución. Los repositorios temáticos albergan colecciones o recursos de una determinada disciplina o área de conocimiento específica; pueden ser creados o dirigidos por instituciones académicas, de investigación y también por organismos gubernamentales. Existen otros tipos de repositorios en función de su alcance, características y tipo de contenidos.

Uno de los mayores precursores y promotores del desarrollo de RI es Clifford Lynch, Director Ejecutivo de la Coalition for Networked Information, quien en el año 2003 proporcionó una de las primeras definiciones: Un Repositorio Institucional es un conjunto de servicios que la universidad ofrece a los miembros de su comunidad para la gestión y para la disseminación de los materiales en forma digital creados por la institución y sus miembros.

En el nivel más básico y fundamental, un RI es un reconocimiento a la vida intelectual y académica de una institución que será representada, documentada y compartida en forma digital; y para ello, una responsabilidad primordial de las universidades es implementar mecanismos para el acceso y preservación de los contenidos (Lynch, 2003). Un Repositorio Institucional “es un archivo digital de la producción intelectual creada por una facultad, un equipo de investigación, y estudiantes de una institución y accesible a los usuarios finales dentro y fuera de la institución, con pocas barreras de acceso o ninguna” (Crow, 2002, p. 17).

Además de los Repositorios Institucionales y temáticos, se identifican otros siete tipos de repositorios: a) Huérfanos, creados para el depósito de trabajos de autores o investigadores que no tienen acceso a otro repositorio institucional o temático; b) de Datos, que almacenan y preservan los datos científicos asociados a un proyecto de investigación; c) Centralizados o especializados, creados por las sociedades científicas, las asociaciones, las entidades gubernamentales o los organismos de financiamiento; d) Multi-institucionales, que integran la producción intelectual de un conjunto de instituciones pertenecientes a una red o consorcio a nivel regional o nacional; e) Integradores o agregadores, son recolectores, agregadores o portales que recolectan los contenidos de varios repositorios institucionales o temáticos; f) de Revistas arbitradas, que almacenan y preservan los títulos de revistas científicas publicadas o financiadas por una determinada institución; g) por tipo de contenidos, son repositorios que almacenan solamente determinado tipo de documentos.

Respecto al tipo de contenidos que pueden albergar los RI, Alonso, Subirats y Martínez (2008) los clasifican en tres tipos: productos científicos, productos institucionales o administrativos y objetos de aprendizaje. Adicionalmente, algunos repositorios institucionales, sobre todo los de España, incluyen acervos antiguos o documentos patrimoniales en sus colecciones.

Es importante aclarar que un RI no es un canal de publicación, sino una vía de comunicación científica, de difusión y visibilidad del conocimiento y debe comprenderse como complementario al proceso de publicación científica formalizado con revisión por pares (Bustos y Fernández, 2008).



Soluciones de software y servicios de metadatos asociados a los repositorios

Existen diferentes aplicaciones de software de código abierto (open source) para crear repositorios institucionales o temáticos, entre los más utilizados se encuentran DSpace, EPrints, Fedora e Islandora. También existen programas comerciales, como Digital Commons, CONTENTdm, DigiTool y EQUILLA Repository.

Para la recuperación del contenido de los repositorios digitales es necesario implementar servicios de metadatos a través de protocolos de intercambio de información. El estándar Dublin Core es uno de los esquemas de metadatos más utilizados que define 15 elementos básicos para describir un documento. También existen estándares para describir recursos más complejos o una agrupación de recursos, como los estándares METS o MODS. Para los objetos de aprendizaje se utilizan el SCORM y el LOM. Para las tesis digitales se utiliza el estándar ETD-MS.

Por lo general, los servicios de metadatos implementan el protocolo estándar de recolección de archivos abiertos, denominado OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting). El protocolo es un servicio Web que permite recuperar la información del repositorio, sus colecciones, los identificadores de los recursos y los metadatos de los recursos.

El proceso de recuperación de metadatos se conoce como recolección o cosecha (*harvesting*). Estos registros se almacenan en una base de datos o en un indexador con una estructura determinada, los cuales pueden ser consultados a través de una interfaz Web.

Algunos de los programas libres disponibles para la creación de integradores o recolectores son DSpace, VuFind, Open Harvester System OHS de PKP y el programa de D-Net del proyecto DRIVER, que es específicamente para infraestructuras complejas.

Existen servicios que validan y certifican los servidores de metadatos de acuerdo al estándar OAI-PMH. Estos servicios se denominan validadores y permiten evaluar y garantizar la interoperabilidad de los repositorios institucionales. Para ello, es necesario utilizar vocabularios controlados o especificaciones de interoperabilidad como DRIVER 2.0 y OpenAIRE, en Europa.

LA RED MEXICANA DE REPOSITARIOS INSTITUCIONALES-REMERI

La Red Mexicana de Repositorios Institucionales es una red federada de repositorios institucionales y temáticos de las Instituciones Mexicanas de Educación Superior y de Investigación, que recolecta e integra su producción científica, académica y documental, para su difusión, visibilidad y acceso abierto. Se puede definir también como un Repositorio agregador de alcance nacional que contiene las referencias (metadatos) de repositorios digitales y revistas de acceso abierto de México.



Antecedentes y desarrollo

El proyecto REMERI surge en principio de la necesidad de contar con una plataforma para integrar los repositorios digitales de las Instituciones Mexicanas de Educación Superior que permitiera su difusión, localización y visualización de manera interoperable a través de interfaces comunes.

Con este propósito, en el 2011 la Biblioteca Virtual de la Universidad Autónoma de San Luis Potosí (UASLP), presenta una iniciativa para desarrollar una Red de Repositorios Institucionales de las instituciones miembros de la Red Abierta de Bibliotecas Digitales (RABID), perteneciente a la comunidad de bibliotecas digitales de la Corporación Universitaria para el Desarrollo de Internet 2 (CUDI).

A partir de esta propuesta, se integra un grupo de trabajo de seis instituciones miembros de RABID, mismas que se constituirían posteriormente como fundadoras de REMERI: Universidad de las Américas Puebla (UDLAP), Instituto Tecnológico de Estudios Superiores de Monterrey (ITESM), Universidad de Guadalajara (UDG), Universidad Autónoma de San Luis Potosí (UASLP), Universidad Autónoma del Estado de Hidalgo (UAEH) y Universidad Autónoma del Estado de México (UAEM).

Por otro lado, en octubre de 2011 se lleva a cabo en la ciudad de México, D.F., la 4^o Reunión de trabajo de miembros del proyecto impulsado por la RedCLARA y financiado por el Banco Interamericano de Desarrollo denominado "*Red Federada de Repositorios Institucionales de Publicaciones Científicas LA-Referencia*". En este proyecto participan representantes de 9 países de Latinoamérica: Brasil, México, Argentina, Chile, Colombia, Perú, Ecuador, Venezuela y El Salvador.

En el marco de esta reunión y ante la necesidad de contar con una plataforma tecnológica que permitiera la creación de un nodo mexicano interoperable con LA-Referencia, los representantes por parte de México en dicho proyecto, el Consejo Nacional de Ciencia y Tecnología (CONACyT) y CUDI, adoptan la iniciativa de REMERI para construir un primer prototipo con base en estándares internacionales predefinidos.

De esta manera, en noviembre del 2011 bajo el liderazgo de la UASLP, el apoyo de CUDI y el financiamiento de CONACyT, el grupo de trabajo de las instituciones fundadoras de REMERI, presentan el proyecto con el propósito de crear una Red federada de repositorios digitales de acceso abierto y la implementación de un primer prototipo nacional para LA-Referencia, designando a la UASLP como el responsable técnico del nodo mexicano en LA-Referencia.

Durante el 2012 se lleva a cabo la primera fase que comprendió el desarrollo de seis componentes estratégicos. Como parte de los resultados, se diseñaron documentos normativos para la operación de la Red, los requisitos técnicos de adhesión y modelos de sostenibilidad financiera; se llevó a cabo un diagnóstico sobre el desarrollo de repositorios institucionales sobre una muestra representativa de 55 instituciones; se organizaron presentaciones y talleres de difusión y capacitación tanto virtuales como presenciales; y finalmente, se implementó la plataforma tecnológica de REMERI y un primer piloto interoperable con LA-Referencia, con 6 Repositorios Institucionales y 3,696 registros indexados.



En abril de 2013, con el apoyo de CUDI y bajo la coordinación general de la UASLP se conforma un grupo técnico de soporte para implementar la segunda fase de REMERI con cuatro objetivos fundamentales: Incorporar a más instituciones y repositorios a la Red y al nodo de LA-Referencia; capacitar y brindar apoyo a las instituciones que no contaban con repositorios estandarizados; consolidar el desarrollo de la plataforma tecnológica; y, formalizar la estructura y administración de la Red para su sostenibilidad operativa y financiera a largo plazo.

Durante este segundo año, a través de estrategias de asesoría y análisis y normalización de metadatos, se logró la integración y recolección de 53 repositorios institucionales y temáticos pertenecientes a 27 Instituciones Mexicanas de Educación Superior con un total de 179,546 documentos. Además, se implementaron mejoras al sitio web y nuevos servicios y herramientas.

Posteriormente, en noviembre de 2013 CUDI convoca a una reunión de miembros fundadores para establecer el modelo de gobernanza de la Red, mismo que es aprobado por su Consejo Administrativo en Febrero de 2014, constituyéndose a REMERI como uno de sus proyectos estratégicos y nueva comunidad de CUDI.

Actualmente (agosto de 2014), REMERI cuenta con 69 repositorios de 37 instituciones para un total 208,695 registros de metadatos de tesis, artículos y libros. De este conjunto, casi 100,000 registros de 30 instituciones están indexadas por LA-Referencia.

Servicios de REMERI

Los servicios que ofrece REMERI son los siguientes:

- Un portal web con noticias, redes sociales, eventos, documentos, informes de incorporación y colecciones, enlaces a servicios, formularios de registro y consultas, directorio de participantes, enlaces a repositorios, servicio de consulta, información de capacitación, material informativo y multimedia.¹
- Un servicio de búsqueda para todas las colecciones con resultados ordenados y filtrados por relevancia, fecha, institución, autor y tipo de documento. Paginación y segmentado de resultados. Ficha informativa y detallada de los recursos con iconografía, enlace al documento, al repositorio y a sus metadatos.²
- Servicio de validación de servidores de metadatos con consulta y validación sintáctica (en base al *Schema*) de la respuesta a los verbos del protocolo *OAI-PMH* con manejo de errores.³
- Formulario de registro para Repositorios Institucionales con información de contacto, descripción del contenido, contactos administrativos y técnicos así como enlaces al repositorio y a su servidor de metadatos. El proceso de registro facilita el proceso de análisis y diagnóstico del repositorio para ser considerado para su incorporación.⁴

¹ <http://www.remeri.org.mx>

² <http://www.remeri.org.mx/portal/REMERI.jsp?busca=tesis>

³ <http://www.remeri.org.mx/portal/valida.html>

⁴ <http://www.remeri.org.mx/portal/formulario-registro.html>

- Una vez cosechado, normalizado e integrado el repositorio, se incorpora al *INDIXE de Repositorios Institucionales* de REMERI, un directorio disponible para consulta desde.⁵
- A su vez el contenido se reporta en el *INDIXE de Producción Científica*, una relación de artículos y tesis de las instituciones de educación superior del país basado en el contenido de sus repositorios.⁶
- Servidores de metadatos institucionales, es un servicio de metadatos de aquellas colecciones y repositorios que no cuentan con los mecanismos para implementar el servicio en sus plataformas.⁷
- Servidor de metadatos de la red, es un servicio de metadatos que permite recuperar todas las colecciones y registros de la red.⁸
- Servidor de metadatos de producción científica estandarizado de acuerdo a los requerimientos de LA-Referencia (basados a su vez en *DRIVER*).⁹

AGREGADOR Y SISTEMA DE CONSULTA INDIXE

INDIXE es un servicio cosechador-agregador (integrador) ideal para repositorios temáticos, regionales o nacionales, desarrollado específicamente para el proyecto REMERI. Es al mismo tiempo un proveedor de datos ya que cuenta con un servicio de metadatos que le permiten a la colección integrarse a otras redes.

Requerimientos técnicos de REMERI

Los requerimientos funcionales definidos para el proyecto REMERI fueron los siguientes: integración en un mismo sistema de diversas colecciones de Repositorios Institucionales; recuperación de información por medio de una interfaz de búsqueda y consulta web; indexación de la colección para mostrar resultados por relevancia; interoperabilidad de la colección utilizando el protocolo OAI-PMH; uso y normalización de los metadatos a estándares y especificaciones Dublin Core, LA-Referencia y DRIVER.

Los requerimientos técnicos definidos para el desarrollo de la plataforma tecnológica consideraron lo siguiente: simplificar la instalación y administración de componentes, facilitar la gestión y administración de los servicios de preferencia con interfaz vía web, permitir procesar y rectificar metadatos recolectados, ofrecer además un servidor de metadatos de la información recolectada de acuerdo a DRIVER.

⁵ <http://www.remeri.org.mx/repositorios>

⁶ <http://www.remeri.org.mx/produccion>.

⁷ http://www.remeri.org.mx/indixe/rest/db/remeri/servicios/unam/tesis/oai_server_tesis_unam.xq?verb=Identify

⁸ http://www.remeri.org.mx/indixe/rest/db/remeri/oai/oai_server.xq?verb=Identify

⁹ http://www.remeri.org.mx/indixe/rest/db/remeri/driver/driver_server.xq?verb=Identify

Para el desarrollo e implementación de este proyecto se consideraron las siguientes alternativas :

- OAIConnect. (www.drupal.org/project/oaiconnect). Permite la recolección de metadatos y los almacena en una estructura estándar en una base de datos relacional. Funciona sobre Drupal, Solr, PHP y MySQL con servidor Apache. Se requerían instalar y configurar tres servicios, no consideraba la transformación de metadatos y no contaba con servidor de metadatos.
- DSpace. (www.dspace.org). Además de gestionar contenidos, permite la recolección de metadatos y los almacena en una base de datos relacional. Funciona con Java sobre Apache Tomcat, Solr, PostgreSQL u Oracle con JSP. Se requerían instalar y configurar tres servicios, no consideraba la transformación de metadatos y en su momento no contaba con el estándar DRIVER para su servidor de metadatos.
- D-Net. (www.d-net.research-infrastructures.eu). Proyecto promovido por la Unión Europea en su iniciativa DRIVER. Permite la agregación de registros y metadatos, transformación y limpieza de metadatos y los almacena en una base de datos no relacional (JSON style). Funciona con Java sobre Jetty, MongoDB, Python. Se requerían instalar y configurar tres servicios, no consideraba la transformación de metadatos.
- VuFind (www.vufind.org) plataforma para la gestión y consulta de registros bibliográficos, incluye recolector de metadatos, almacena la información en Solr y funciona con PHP y MySQL con servidor Apache. Se requerían instalar y configurar tres servicios, no consideraba la transformación de metadatos y en su momento no consideraba el estándar DRIVER para su servidor de metadatos.
- Desarrollo propio. Como alternativa se contaba con la posibilidad de desarrollar la solución basándose en los requerimientos y particularidades del proyecto.

Se consideraron estas opciones y se decidió por el desarrollo propio con la plataforma eXist XML-DB de código abierto. Dentro de las consideraciones para su elección se encontraron: que es una plataforma de desarrollo todo en uno (base de datos, administración web y servidor web); el uso de una misma tecnología (XML) para almacenamiento de la información, procesamiento y consulta; indexación del texto de la colección para consultas de documentos por relevancia; el uso del mismo lenguaje para todo el desarrollo (XQuery) que no requiere compilación ya que es interpretado; interfaz de gestión y administración web, que no requiere el acceso a consola del servidor; instalación una sola vez y un único servicio; plataforma web robusta y probada (Java-Tomcat).

Además, en un primer análisis de los repositorios disponibles en México para la construcción del prototipo, se encontraron servidores con metadatos no estandarizados, inconsistentes, faltantes o sin estructura, con implementaciones parciales o con modificaciones del protocolo OAI-PMH; adicionalmente, una diversidad de recursos, formatos y áreas de conocimiento, por lo que en su momento ningún repositorio analizado cumplía con las especificaciones de DRIVER.

Arquitectura y funcionalidad del Sistema INDIXE

Para la cosecha, indexación y consulta de los repositorios agregados a REMERI, se desarrolló una herramienta propia denominada INDIXE basado en tecnologías XML (XQuery, XPath, XSLT). La base de datos utilizada (eXist) almacena y procesa los metadatos en el formato XML que es el formato nativo en la que se encuentran los metadatos (Dublin Core). La programación de tareas, procesos, servicios y consultas se realizaron en el lenguaje XQuery especializado para información semiestructurada. La base de datos cuenta con indexación de la colección con Lucene con un método de espacios vectoriales y booleano. Se consiguió con esto un desarrollo escalable, de alto desempeño, eficiente, con código compacto y multiplataforma. La arquitectura del sistema INDIXE se muestra en la figura 1.

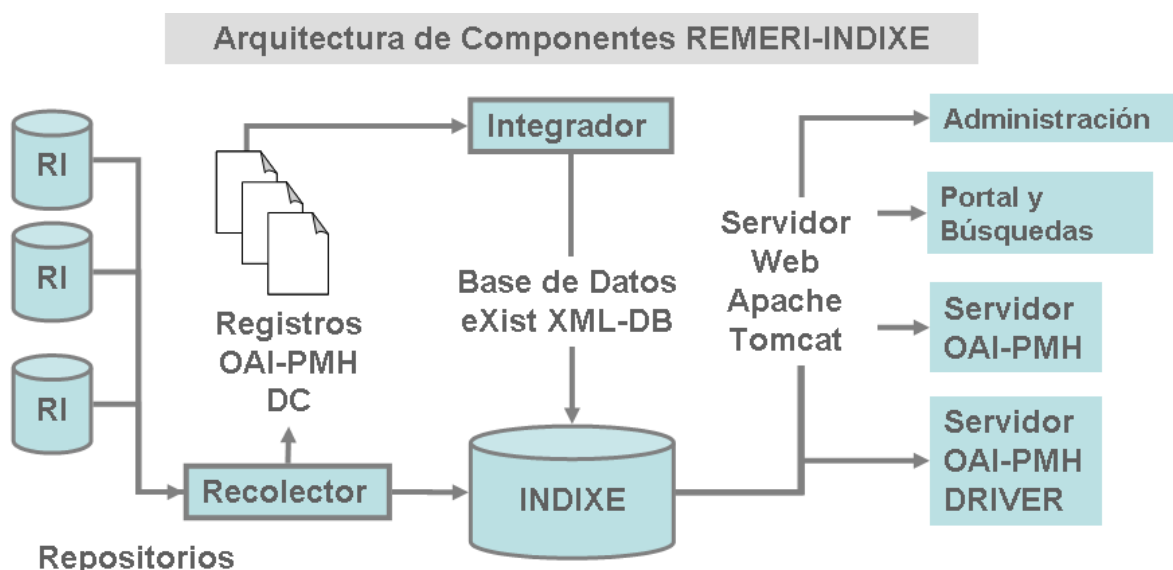


Figura 1. Diagrama de Componentes

Dentro de las aplicaciones o herramientas desarrolladas en el proyecto se encuentran implementadas en el lenguaje XQuery y son consultadas vía web:

- Un validador (sintáctico) para servidores de metadatos
- Un servicio de recolección (cosecha)
- Un servicio integrador (con normalización y estructura de metadatos)
- Un servicio de consulta y recuperación de información (ordenado por relevancia, fecha, institución, tipo o autor)
- Un servidor de metadatos en el estándar OAI-PMH,
- Un servicio integrador para LA-Referencia (DRIVER) con normalización de tipo de documentos y la incorporación del campo de institución (instname) y repositorio (reponame)
- Un servidor de metadatos en el estándar OAI-PMH para LA-Referencia (DRIVER) con especificaciones técnicas requeridas con tipos de documentos y tamaño de la consulta (token).



La infraestructura tecnológica sobre la que opera la plataforma de software consiste en un servidor Dell PowerEdge R720xd con dos procesadores, 64 GB en RAM y 3.5 TB en disco con sistema operativo Linux (OpenSUSE 12.1 amd-64). La plataforma de desarrollo (JSP), el servidor de web y de aplicaciones (Tomcat 6.0) junto la base de datos (eXist 1.4.2) funcionan con Java (OpenJDK 1.6).

INTEROPERABILIDAD DE REMERI CON REDES FEDERADAS DE REPOSITARIOS INSTITUCIONALES DE PUBLICACIONES CIENTÍFICAS

En REMERI pueden participar instituciones de educación superior públicas y privadas, centros de investigación, organizaciones sin fines de lucro y, en general de instituciones de los sectores público y privado en México. Todo miembro de la Red debe configurar su repositorio de forma que permita compartir información de forma homologada, utilizar un estándar de metadatos preferentemente Dublin Core y se hace una fuerte recomendación de cumplimiento con el protocolo OAI-PMH y las directrices DRIVER 2.0. Los metadatos requisito que todo miembro de REMERI debe asegurarse de configurar en su repositorio, son los siguientes:

- dc:title
- dc:identifier
- dc:creator
- dc:type
- dc:date

Los metadatos que en caso de no localizarse pueden ser procesados y asignados de forma automatizada a través del sistema de recolección de metadatos son los siguientes:

- dc:rights
- dc:publisher
- dc:language
- dc:format
- dc:source

Los metadatos que no son requisito, sin embargo se hace una fuerte recomendación que se llenen de forma correcta y se provean, son los siguientes:

- dc:description
- dc:subject
- dc:contributor



El servidor de metadatos del Repositorio Institucional debe aprobar el proceso de validación (sintáctica) y una vez analizado y evaluado su contenido, se establece el proceso de integración. El proceso de recolección requiere la completa implementación de los verbos de OAI-PMH para recuperar los identificadores (ListIdentifiers) y cada registro de manera particular (GetRecord). Las recolecciones se realizan completas cada vez, para que en el caso de hacer ajustes en el proceso de transformación, la colección resulte consistente. El proceso de transformación consiste en mapear, normalizar y estandarizar los metadatos (de acuerdo a los casos presentados anteriormente) y al mismo tiempo, genera la versión de los registros para su consulta e incorporación al servidor de metadatos.

Interoperabilidad con LA- REFERENCIA

Para el caso de las colecciones definidas en LA-Referencia, las tesis de maestría, doctorado, artículos y reportes de investigación (únicos documentos establecidos por al Red para ser cosechados), se cuenta por cada repositorio con un proceso de transformación de acuerdo a los requerimientos de DRIVER para el uso de vocabularios, términos y elementos, a su vez, se incorpora la colección del servidor de metadatos DRIVER para ser incorporados al servidor de metadatos consultado por LA-Referencia.

La experiencia en este proyecto nos ha permitido identificar problemáticas comunes en los diferentes tipos de repositorios, como las siguientes:

dc:identifier

Es común en el caso del software DSpace, exponer el uso del identificador con un handle cuando este no se encuentra activo para la institución y el servidor, también se ha encontrado el uso del IP del servidor o el término "localhost". En muchos casos los administradores de repositorios no están al tanto de esto y pueden contar con una gran cantidad de registros expuestos de esta manera. Esto se puede corregir en el proceso de integración con INDIXE pero el repositorio original no es consistente para otros proyectos. Es recomendable el acceso al documento directamente, no a páginas de presentación o intermedias, de manera que se agilice la consulta al registro.

dc:type

Es común encontrar el tipo "otros" o registros sin tipos. Es mejor definir en el repositorio a detalle los tipos de materiales para su correcta consulta e incorporación. Para repositorios temáticos es posible asignar los tipos de manera automática. En el caso del proyecto REMERI y LA-Referencia es recomendable agregar o especificar el tipo de tesis, "Tesis de Maestría" o "Tesis de Doctorado" al tipo genérico "Tesis". En el caso de repositorios con conjuntos (DSpace) esto se puede resolver de manera automatizada en el proceso de integración pero hay que analizar caso por caso. Lo más recomendable es utilizar los vocabularios de DRIVER para tipos de datos.

dc:date

Es común encontrar más de una ocurrencia para las fechas en el caso de plataformas de gestión (se incluye la fecha de registro y la fecha de última actualización). La fecha de publicación generalmente se encuentra en la misma posición y es posible recuperarla con INDIXE de manera automatizada. Existen registros sin fecha o con la fecha en el formato no estándar, también se procesa en la integración nuevamente caso por caso. Lo mas recomendable es mostrar una sola fecha.

dc:publisher

Casi ningún proveedor (por default) proporciona el nombre de la Institución y el repositorio de dónde se está obteniendo la información. Se agrega esta información como metadato a cada registro, no se debe de obviar ya que los metadatos son procesados automáticamente e integrados con otros, el nombre del identificador o URL no siempre es significativo y menos aún cuando se utiliza un IP. Esto también se procesa en la integración en cada caso requerido.

En el caso de LA-Referencia (www.lareferencia.info), REMERI formó parte de las pruebas técnicas desde octubre del 2012. Hasta el momento, México es la red nacional que incorpora la mayor cantidad de registros en idioma Español en el proyecto con un total de 95,194 (en agosto de 2014) provenientes de un total de 55 repositorios institucionales de 31 instituciones, manteniendo un crecimiento continuo y actualizado. Las estadísticas de recolección de México en la referencia por fecha se pueden ver en la figura 2.



Figura 2. Estadísticas de cosecha de materiales de México en LA-Referencia disponible en: <http://www.lareferencia.info/vufind/Laref/Cosecha?iso=MX>



REMEDI fue la primera red nacional en cumplir con lineamientos técnicos específicos de LA-Referencia, incluyendo los términos "instname" y "reponame" y requerimientos para DRIVER. El porcentaje de transformación (adecuación) de metadatos por parte de LA-Referencia es de 0.1 %, siendo el nodo nacional con el menor número de transformaciones y que tiene el mayor porcentaje de aceptación de registros, equivalente al 99.2 % en agosto 2014.

RESULTADOS Y CONCLUSIONES

Después de más de dos años de trabajo, REMEDI cuenta en este momento (agosto del 2014) con la incorporación de 69 repositorios de 37 Instituciones Mexicanas para un total de 208,695 documentos incluyendo artículos (24.7%) , tesis de licenciatura (15.6%), tesis de maestría (21.9%), tesis de doctorado(6%) e imágenes (25%). En menor porcentaje se encuentran los videos, objetos de aprendizaje, trabajos recepcionales, capítulos de libros y libros.

Hasta el momento con INDIXE se han desarrollado 19 servidores de metadatos institucionales, incorporando 25 repositorios universitarios de REDALYC y mantenido el servidor de metadatos del proyecto interoperable y registrado en los directorios OpenDOAR y ROAR y con LA-Referencia. Además, se están desarrollando pruebas con INDIXE para gestionar redes de repositorios de producción cultural, revistas científicas y de tesis digitales.

Por otro lado, entre los principales retos a nivel nacional para el desarrollo y consolidación de repositorios digitales de acceso abierto, podemos mencionar los siguientes: La capacitación al personal responsable de la gestión de los Repositorios en tres competencias básicas: tecnológicas, informacionales y de comunicación; la estandarización y normalización de metadatos conforme a directrices internacionales de interoperabilidad; y, el establecimiento de políticas y mandatos de Acceso Abierto en las Instituciones Mexicanas de Educación Superior.

Finalmente, el proyecto REMEDI que inicio con el desarrollo de un plan piloto, es ahora una Red de alcance nacional, que cumple con los estándares y directrices internacionales de interoperabilidad, siendo el primer proyecto mexicano en su tipo que además, está alineado a la estrategia nacional que está siendo coordinada por el CONACyT para el desarrollo del Repositorio Nacional de Producción Científica e Innovación, en el marco de la nueva Ley de Acceso Abierto de México.



REFERENCIAS BIBLIOGRÁFICAS

Abadal, E. (2012). *Acceso abierto a la ciencia*. Editorial UOC. Recuperado el 18 de febrero de 2014, desde:
<http://diposit.ub.edu/dspace/bitstream/2445/24542/1/262142.pdf>

Alonso, J., Subirats I. y Martínez, M. L. (2008). Informe APEI sobre acceso abierto. Gijón: Asociación Profesional de Especialistas en Información. Recuperado el 9 de diciembre de 2013, desde: <http://eprints.rclis.org/12507/>

Bustos G. A., y Fernández P. A. (2008). *Directrices para la creación de repositorios institucionales en universidades y organizaciones de educación superior*. Recuperado el 13 de diciembre de 2013, desde:
<http://repository.urosario.edu.co/handle/10336/223>

Budapest Open Access Initiative (2002). Recuperado 28 de enero de 2014, desde:
<http://biblioteca.upc.es/rebiun/BOAI.pdf>

Crow, R. (2002). The Case for Institutional Repositories: A SPARC Position Paper. Recuperado el 3 de diciembre de 2013, desde:
http://works.bepress.com/ir_research/7

Lynch, C. (2003). Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age ARL: A Bimonthly Report, no. 226 (February 2003). Recuperado el 30 de noviembre de 2013, desde:
<http://www.arl.org/storage/documents/publications/arl-br-226.pdf>

Suber (2012). *Open Access*. Cambridge, Massachusetts: The MIT Press Essential Knowledge Series. Londres, Inglaterra.



CONFIGURAÇÃO E DESENVOLVIMENTO DO ACERVO DE RECURSOS EDUCACIONAIS EM SAÚDE DA UNA-SUS

Onivaldo Rosa Junior¹, Aline dos Santos Jacob²

¹Engenheiro de Software da Secretaria Executiva da Universidade Aberta do SUS, Mestre em Engenharia de Produção pela Universidade Federal de Santa Catarina.

E-mail: onivaldo@unasus.gov.br.

²Bolsista da Universidade Aberta do SUS, Pós-graduanda em Repositórios Digitais na FIOCRUZ/RJ

E-mail: alinejacob@unasus.gov.br.

RESUMO

O Acervo de Recurso Educacional em Saúde - ARES é o repositório educacional digital da Universidade Aberta do SUS – UNA-SUS onde são disponibilizados recursos utilizados nas ações de ensino e aprendizagem para trabalhadores do SUS. O repositório foi elaborado utilizando o software *DSpace*, específico para gerenciamento de conteúdo digital. Para adequação deste software as necessidades da UNA-SUS foram necessários alguns procedimentos técnicos de adaptação, principalmente no que tange ao suporte a metadados específicos relacionados a avaliação do recurso assim como metadados de interoperabilidade que definem o público-alvo (utilizando o Código Brasileiro de Ocupações - CBO) e descritivo utilizando Descritores em Ciências da Saúde - DeCS, ambos integrados por meio de webservices. Este artigo tem como objetivo descrever esses procedimentos, indicando as principais adequações e configurações específicas aplicadas a versão 1.8.0 do *DSpace*, realizadas com interferência mínima no código fonte original e concordância com as diretrizes publicadas na política de desenvolvimento do acervo. Como resultados, são detalhados os esquemas de codificação utilizados para cada alteração feita no *DSpace*, desde alterações relacionadas a submissão, gestão de usuários e integração com a Federação de Autenticação UNA-SUS (que utiliza o padrão SAML 2.0). As adequações utilizadas estão permitindo a migração rápida e descomplicada para versões mais atuais.

Palavras-chave: Metodologias Computacionais, Gestão da Informação em Saúde, Integração de Sistemas.



ABSTRACT

The ARES (Acervo de Recurso Educacional em Saúde) is the Collection of Educational Resource in Health of Universidade Aberta do SUS – UNA-SUS, available resources are being used in the actions of teaching and learning focusing on professional of health. The collection was compiled using the DSpace software, specific to digital content management. Suitability of this software for the needs of the UNA-SUS took some technical adaptation procedures, especially with regard to support specific metadata related to resource assessment as well as interoperability of metadata defining the target audience using the Brazilian Code of Occupations (CBO) and using descriptive Descriptors in Health Sciences - DeCS, both integrated through webservice. This article aims to describe these procedures, indicating the main adaptations and specific settings applied to version 1.8.0 of DSpace, performed with minimal interference in the original source code and compliance with the guidelines published in the collection development policy. The results are detailed encoding schemes used for each change made to DSpace from changes related to submission, user management and integration with the Federation Authentication UNA-SUS (which uses the SAML 2.0 standard). The adjustments used allowing for quick and uncomplicated migration to more current versions.

Keywords: Computing Methodologies, Information Management in Health, Systems Integration.

INTRODUÇÃO

O Acervo de Recurso Educacional em Saúde - ARES é o repositório educacional de recursos digitais em saúde da Universidade Aberta do SUS – UNA-SUS. É formado pelo conjunto de recursos educacionais, sejam objetos de aprendizagem com conteúdo, objetivo e avaliação, ou unidades educacionais menores selecionadas e disponibilizadas por vias de comunicação da UNA-SUS para subsidiar ações de educação em saúde em todo país. Com o ARES, espera-se beneficiar os trabalhadores do SUS oferecendo acesso amplo a conteúdos em saúde que possibilitam oportunidades continuadas de aprendizado (BRASIL, 2010).

O propósito do ARES é preservar e tornar pública a produção de conhecimento legada dos processos educacionais voltados para trabalhadores da saúde em todo país. Diversas atividades educativas para profissionais de saúde são financiadas em todo país pelos poderes públicos municipais, estaduais e federal sem que exista um mecanismo nacional para possibilitar o registro e compartilhamento dessas experiências. O Sistema UNA-SUS visa possibilitar a integração entre essas atividades, por meio dos seus elementos constitutivos: a Rede UNA-SUS – formada por Instituições de Ensino Superior reconhecidas pelo MEC –, a Plataforma Arouca – um sistema de informações nacional de históricos dos profissionais de saúde –, e o ARES. (BRASIL, 2012).



Para que os diversos tipos de recursos que integram o acervo estejam disponíveis aos usuários, o repositório foi organizado em coleções com características, tratamento e organização específicas. Essa organização foi possível após a implantação de requisitos tecnológicos em ferramenta própria, escolhida após estudos e avaliação técnica. Optou-se, assim, por utilizar como base a plataforma de software *Dspace*, principalmente pelas seguintes características:

- Possuir código fonte aberto, sobre condições de licença própria;
- Possuir uma comunidade estabelecida de desenvolvedores;
- Personalizável e modular;
- Ser utilizado amplamente no mundo, por instituições educacionais, governamentais e privadas;
- Capacidade de gerenciamento de diversos tipos de conteúdo digital.

Diante disso, o presente artigo tem como objetivo descrever os procedimentos técnicos de adaptação da ferramenta *Dspace*, as alterações no código e configurações específicas necessárias para implantação do Acervo de Recursos Educacionais em Saúde da UNA-SUS.

Procedimentos Metodológicos

Para implantação do ARES foi preciso considerar os requisitos tecnológicos fundamentais para garantir sua interoperabilidade com outros sistemas, definidos por meio de estudos, além de necessidades preestabelecidas, tais quais:

- Utilizar como base o código fonte da ferramenta *Dspace* em sua versão 1.8.0;
- Realizar as alterações de modo a minimizar ao máximo a interferência no código fonte original, permitindo a migração destas alterações para versões futuras do *Dspace*;
- Uso obrigatório da autenticação federativa da UNA-SUS (padrão SAML versão 2 (OASIS 2012), que define a obrigatoriedade de nome completo, CPF e e-mail para todos os usuários;
- A configuração e customização deveriam respeitar as diretrizes publicadas no documento “Política do Acervo de Recursos Educacionais em Saúde (ARES)”, com ênfase no seu Anexo C (Metadados). (BRASIL, 2011).
- Integração com a CBO - Classificação Brasileira de Ocupações (BRASIL, 2002) e Integração com o DeCS - Descritores em Ciências da Saúde (Biblioteca Virtual em Saúde, 2014).



CONTEXTO DE APLICAÇÃO

Sistema Operacional: Centos 5.7

Ferramentas instaladas: JAVA SDK 1.7, ANT 1.8.2, MAVEN 3.0.3 (Java SDK, ANT e MAVEN são necessários para o processo de customização do código fonte), configurações específicas utilizadas: JAVA_OPTS="-Xmx512M -Xms64M -Dfile.encoding=UTF-8" MAVEN_OPTS="-Xmx512m -Xms256m"

Servidores WEB: Apache 2.2.3 com SSL e Módulo Shibboleth 2.4.3 , Tomcat 7.0.8

Interface escolhida: XMLUI, com interface Discovery ativada (Não foram realizadas alterações em outras interfaces não utilizadas pelo ARES).

Gerenciados de Banco de Dados: PostgreSQL 9.1.2

Módulos de Interoperabilidade: OAI-PMH, OPENSEARCH (LNI, SWORDS em estudos)

Módulo Authority para uso do vocabulário controlado CBO, utiliza um Webservice criado pela UNA-SUS

Caracterização do Webservice CBO (Cadastro Brasileiro de Ocupações) criado pela SE/UNA-SUS para uso no ARES, suprimindo a falta de um Webservice Oficial do Ministério do Trabalho. A base de dados foi obtida em: www.mteco.gov.br/cbsite
Link de acesso: http://sistemas.unasus.gov.br/ws_cbo/cbo.php

Acesso tipo 1)

Complemento parâmetro: words=#ARES ou sem parâmetros

Retorno: Lista de registros de ocupações profissionais de saúde de nível superior para efeito de classificação ARES, detalhes na página do webservice.

Acesso tipo 2)

Complemento parâmetro: words=palavra_base

Retorno: Lista de registros de ocupações profissionais que tenham a palavra_base no nome seja na designação oficial ou em sinônimos.

Acesso tipo 3)

Complemento parâmetro: cbo=codigo_cbo

Retorno: Registro de ocupação profissional que corresponda ao codigo_cbo.



Para permitir o uso de famílias de CBOs, adotamos o padrão de substituir os dois dígitos finais com XX, para a designação de famílias em todos os processos do Webservice e no ARES.

Exemplo: http://sistemas.unasus.gov.br/ws_cbo/cbo.php?words=parteira

```
<unasus_cbo version="1.0" date="20120223 212622" query="parteira">
  <cbo_response>
    <cbo>515115</cbo>
    <descricao>Parteira leiga</descricao>
    <sinonimos>Parteira, Parteira prática</sinonimos>
  </cbo_response>
  <cbo_response>
    <cbo>322230</cbo>
    <descricao>Auxiliar de enfermagem</descricao>
    <sinonimos>Auxiliar de parteira</sinonimos>
  </cbo_response>
  <cbo_response>
    <cbo>223545</cbo>
    <descricao>Enfermeiro obstétrico</descricao>
    <sinonimos>Enfermeira parteira</sinonimos>
  </cbo_response>
</unasus_cbo>
```

A codificação do módulo para uso deste webservice foi realizada com o desenvolvimento de uma classe CBO, uma classe de apoio Base64, e uma classe principal denominada CBOAuthority que por sua vez implementa uma interface do tipo ChoiceAuthority. Esta interface define o padrão de funcionamento das classes do tipo Authority, que permitem definir de forma customizada fontes de dados para campos de metadados que serão controlados por uma autoridade externa ao repositório. Neste caso um webservice com as características citadas anteriormente.

Em termos de interface obteve-se o resultado apresentado na Figura 1, e o dado armazenado pode ser observado na Figura 2.

Figura 1 - Módulo CBO - Interface do usuário

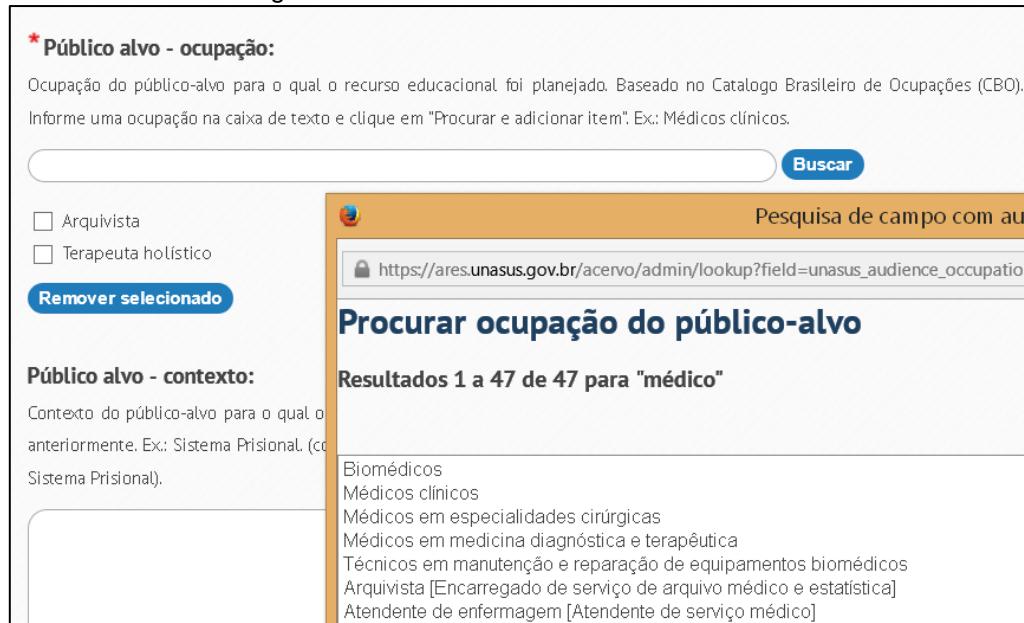
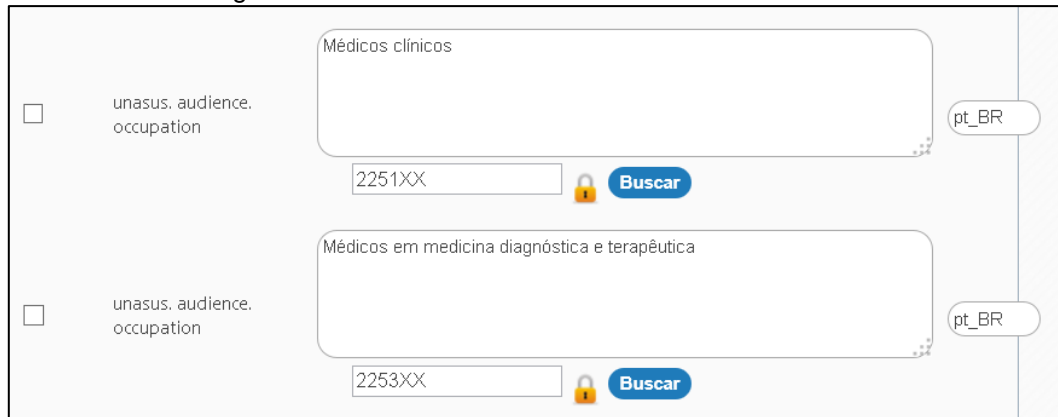


Figura 2 - Módulo CBO - Armazenamento do metadado



Módulo Authority para uso do vocabulário controlado DeCS, utiliza um Webservice criado pela Bireme

Este módulo faz uso intensivo do webservice oferecido pela Bireme, descrito em:
http://wiki.reddes.bvsalud.org/index.php/Serviços_DeCS

Utilizou-se basicamente duas formas de acesso ao webservice DeCS. Uma de consulta por busca de palavra:
http://decs.bvsalud.org/cgi-bin/mx/cgi=@vmx/decs/?words=Macaca_mulatta

E outra com seu complemento reverso, busca por código:

http://decs.bvsalud.org/cgi-bin/mx/cgi=@vmx/decs/?tree_id=B01.050.150.900.649.801.400.112.199.120.510.550&lang=pt

A codificação do módulo para uso deste webservice foi realizada de forma similar ao módulo CBO, neste caso utilizando uma classe principal denominada DeCSAuthority.

Em termos de interface obteve-se o resultado apresentado na Figura 3, e o dado armazenado pode ser observado na Figura 4.

Figura 3 - Módulo DeCS - Interface do usuário

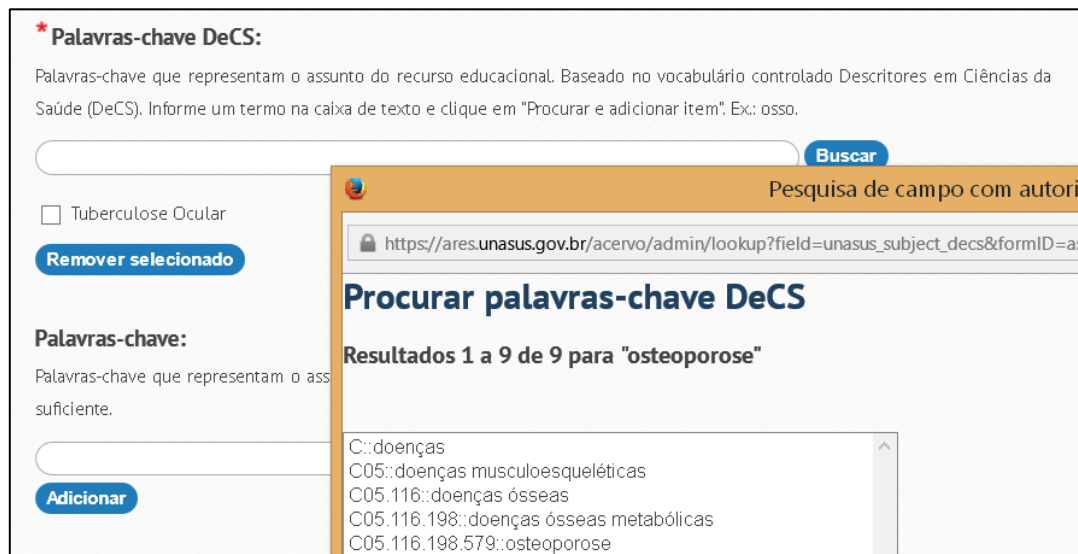


Figura 4 - Módulo DeCS - Armazenamento do metadados

Novo tipo de entrada para atender as necessidades do metadado complexo “Avaliação”

Nas definições de metadados adotadas no ARES, temos três metadados associados a avaliação:

unarus.avaliation.content, utilizado para indicar uma avaliação técnico-científica do recurso educacional.

unarus.avaliation.relevance, utilizado para indicar uma avaliação de pertinência do recurso educacional.

unarus.avaliation.mediaPedagogical, utilizado para indicar uma avaliação em relação a questão midiático-pedagógica do recurso educacional.

Cada um deles indica intrinsecamente três informações relacionadas a respectiva avaliação (Ano que a mesma foi realizada, O Responsável por esta avaliação, seja uma pessoa, grupo, ou instituição e o tipo de avaliação realizada). Como uma característica do DSPACE é a utilização de metadados planos, estas informações não poderiam ser catalogadas de forma hierárquica. Para resolver esta questão no ARES, formalizamos a definição do campo com um campo texto tradicional no estilo dc, com a seguinte marcação para armazenamento: ano::responsável::tipo

Sendo o ano representado por quatro dígitos e o tipo representado pelo dicionário controlado: expert => Perito e user => Usuário

O DSPACE não fornece por padrão nenhum tipo de campo de formulário de entrada de dados que fosse adequado a esta estrutura, portanto, um novo tipo teve que ser desenvolvido. A Figura 5 apresenta este tipo de campo na interface atual do ARES, na Figura 6 é observa-se como é armazenado o metadados.

Para esta modificação foi necessário criar uma classe nova chamada DCUNASUSAvaliation, e uma série de alterações em códigos do core do DSPACE para acomodação deste novo tipo. Nenhuma alteração de banco de dados foi necessária.

Figura 5 - Metadados de Avaliação - Interface com o usuário

*** Avaliação técnico-científica:**
 Avaliação técnico-científica do recurso educacional.

2013 Secretaria de Saúde de Pernambuco Avaliação por perito **Adicionar**
 Ano Responsável Tipo

2014::Maria da Penha Silveira Martins::Avaliação por perito

Remover selecionado

*** Avaliação de pertinência:**
 Avaliação de pertinência do recurso educacional.

... Avaliação por perito **Adicionar**

Figura 6 - Metadado de Avaliação - Armazenamento do metadados

 unasus. evaluation. content 2013::Felipe Teixeira de Melo Freitas::expert

 unasus. evaluation. content 2013::Paulo Afonso Martins Abati::expert

Conforme apresentado os módulos para dicionário CBO e DeCS foram desenvolvidos dentro das especificações do módulo Authority do próprio DSPACE, permitindo a evolução destes módulos nas versões posteriores da ferramenta.

O Tipo de entrada adicional para os metadados de avaliação foi igualmente desenvolvido dentro dos parâmetros do DSPACE, se integrando perfeitamente ao código original.

Módulo de Autenticação Federação UNA-SUS

As customizações envolvidas no modulo de autenticação Shibboleth, e no módulo de perfil de usuários (ePerson) foram relacionadas ao uso das credenciais específicas desta federação: Nome completo, e-mail e CPF. (Figura 7)

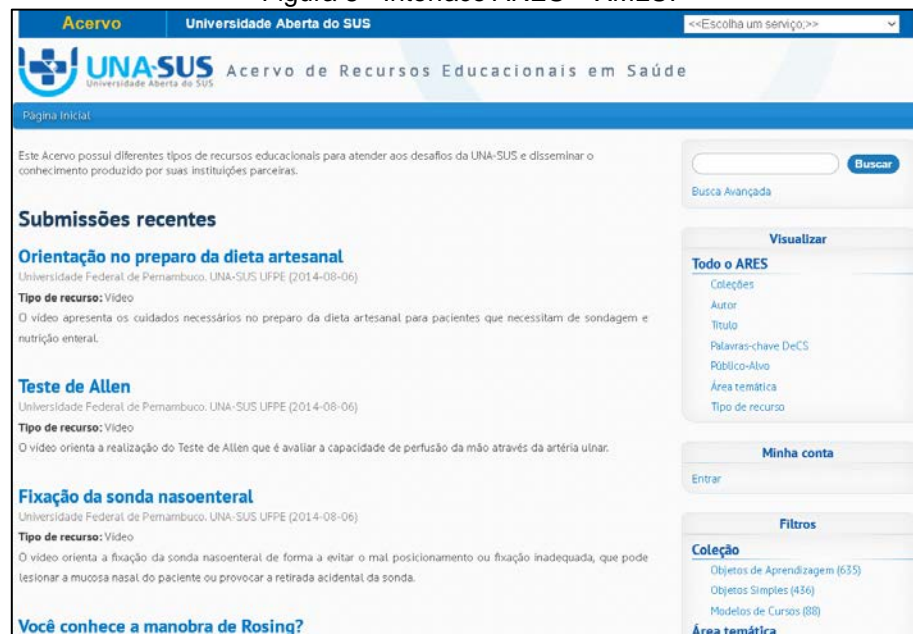
Os dados foram mapeados de modo a reaproveitar o campo netid da tabela eperson do DSPACE para armazenar o CPF, garantindo a integridade e unicidade do cadastro de usuários. Como a federação de autenticação trabalha com o nome completo, o armazenamento do mesmo foi dividido automaticamente considerando apenas uma palavra para o campo firstname e o restante armazenado no campo lastname.

Figura 7 - Interface de login da Federação de Autenticação UNA-SUS

Tema ARES SE/UNA-SUS

Criação de um tema DSPACE UNA-SUS, Figura 8, baseado no tema padrão Mirage (XMLUI) desenvolvido utilizando CSS e XSL, com internacionalização i18n; Algumas customizações no código Java foram realizadas para ajustar os “aspects” (códigos responsáveis pela geração do XML base a ser transformado pelo tema)

Figura 8 - Interface ARES – XMLUI



Customizações no processo de submissão

No processo de submissão foram feitas customizações para obter-se nomenclatura adequada nos passos de submissão e ajustes nas perguntas iniciais (Figura 9).

Figura 9 - Customização das perguntas iniciais e nomenclatura das etapas de descrição

Submissão de recurso

Questões iniciais → Descrição geral → **Descrição educacional** → Descrição temática → Direitos autorais →
 Arquivos → Revisão → Licença → Submissão finalizada

Questões iniciais

Múltiplos títulos:
 O recurso possui mais de um título? *Ex.: título principal em português e título alternativo traduzido para o inglês.*
 Sim

Publicado:
 O recurso foi publicado ou distribuído anteriormente?
 Sim

Recurso Restrito:
 O recurso é de acesso restrito, não podendo ser acessível ao público em geral?
 Sim

Nota importante: Caso o checkbox acima não esteja marcado as configurações de restrição de acesso serão removida do recurso

Salvar e Sair **Próximo >**

Para realizar a customização necessária para adequação do passo “Perguntas Iniciais do processo de submissão; Com a adição da pergunta “O recurso é de acesso restrito, não podendo ser acessível ao público em geral?”, foi necessário adicionar um campo chamado `restricted_access` do tipo booleano nas tabelas `workflowitem` e `workspaceitem` do DSPACE.

Exceções ao uso do Override de Código

Os módulos DeCS e CBO apresentaram problemas quando da utilização do comando de indexação `dspace index-init` sendo isto solucionado se os códigos em `dspace-1.8.0/dspace/modules/xmlui/src/main/java/org/dspace/content/authority` forem também copiados para a pasta do código original do DSPACE `/home/ares/dspace-1.8.0/dspace-api/src/main/java/org/dspace/content/authority`.

Esta não conformidade não prejudica a adequação geral, já que se trata de módulos adicionais, não causando conflito com o código original.



Os ajustes de banco de dados (somente foram alterados os códigos para PostgreSQL, adequações para uso com outros bancos não foram realizadas) devem ser incluídos manualmente após a primeira instalação e antes do primeiro acesso, ou alterar os arquivos em: dspace-1.8.0/dspace/etc/postgres, antes do processo de compilação/instalação.

Outros ajustes menores foram realizados por meio de configurações disponibilizadas pelo próprio DSPACE, assim como a tradução para elementos novos, adicionados pelas expansões desenvolvidas.

CONSIDERAÇÕES FINAIS

A configuração e adaptação da ferramenta respeitando ao máximo seus códigos originais possibilitaram a migração rápida e descomplicada para versões mais atualizadas, a utilização de conceitos intrínsecos os DSPACE permitiram uma integração facilitada com CBO e DeCS, por meio da utilização de extensões disponibilizadas pelos desenvolvedores. Alguns ajustes necessitaram mudanças de códigos considerados do núcleo do sistema, o que nos leva a conclusão que o DSPACE ainda necessita de desenvolvimentos adicionais, e que o processo de customização e integração ainda carece de mecanismos mais amigáveis.

Os procedimentos técnicos de adaptação da ferramenta DSpace para o Acervo de Recursos Educacionais em Saúde (ARES) da UNA-SUS geraram ainda outras necessidades:

- Ajustes do código do ARES para sincronismo com versão 1.8.1 do DSPACE;
- Configurações avançadas utilizando a indexação Apache Lucene;
- Estudo e implementação de relatórios avançados utilizando Apache Lucene;
- Configuração de acesso restrito a determinados recursos sensíveis a Médicos e/ou outras categorias profissionais;
- Sistema de busca avançada usando Apache Lucene e/ou opensearch;
- Configuração avançada da interface OAI-PMH e padrões de acesso.

Atualmente, a SE/UNA-SUS está iniciando o processo de atualização do ARES para aversão mais atual do DSPACE, e já visualizamos melhoras em alguns aspectos de integração e customização, e comprovamos que a abordagem citada neste artigo será essencial neste processo de evolução contínua.



REFERÊNCIAS

Brasil. Decreto Presidencial nº 7.385, de 8 de dezembro de 2010. Institui o Sistema Universidade Aberta do Sistema Único de Saúde - UNA-SUS, e dá outras providências. Diário Oficial da União, 09 dez 2010; Seção 1; 234:1.

BRASIL. Ministério da Saúde. Secretaria de Gestão do Trabalho e da Educação na Saúde. Universidade Aberta do Sistema Único de Saúde. **Política do acervo de recursos educacionais em saúde**. Responsabilidade técnica: Fernanda de Souza Monteiro e Aline Santos Jacob. – Brasília: UNA-SUS, 2011.

BRASIL. Ministério do Trabalho e Emprego. Classificação Brasileira de Ocupações; 2002. Disponível em: <http://www.mteco.gov.br/cbsite/pages/informacoesGerais.jsf>.

Oasis. OASIS Approved Errata. SAML Version 2.0 Errata 05. 01 May 2012. Disponível em: <http://docs.oasis-open.org/security/saml/v2.0/errata05/os/saml-v2.0-errata05-os.html>.

Biblioteca Virtual em Saúde. Descritores em Ciências da Saúde – DeCS. Disponível em: <http://decs.bvs.br/>.



A MIGRAÇÃO DA BIBLIOTECA DIGITAL DE TESES E DISSERTAÇÕES DA UNIVERSIDADE FEDERAL DE GOIÁS (BDTD/UFG): TÓPICOS A SEREM CONSIDERADOS

Carla Lopes Ferreira¹, Cláudia Oliveira de Moura Bueno², Maria Silvério da Silveira³

¹Bibliotecária, bolsista de pesquisa na Gerência de Informação Digital e Inovação (GIDI) da Universidade Federal de Goiás (UFG), Goiânia, GO, Brasil. E-mail: carlaferreira66@gmail.com

¹Bibliotecária na Gerência de Informação Digital e Inovação (GIDI) da Universidade Federal de Goiás (UFG), Goiânia, GO, Brasil. Doutoranda da Universidad de Murcia – ES. E-mail: claudia_bueno@ufg.br

³Bibliotecária, Diretora do Sistema de Bibliotecas da Universidade Federal de Goiás (UFG), Goiânia, GO, Brasil. E-mail: msilverioufg@gmail.com

RESUMO

Este trabalho apresenta dois aspectos da Biblioteca Digital de Teses e Dissertações (BDTD) da Universidade Federal de Goiás (UFG). O primeiro, descrevendo sua criação e as dificuldades impostas na formalização do processo de depósito na instituição, assim como as estratégias para solucioná-las. Os resultados demonstram que a BDTD/UFG possui mais de 2 mil Teses e Dissertações Eletrônicas (TDEs) no *software* TEDE, que são de grande importância para os Programas de Pós-Graduação e para a comunidade universitária como um todo, assim como na promoção do Acesso Aberto à informação técnico-científica brasileira. São demonstrados ainda, dados estatísticos durante o período de funcionamento do TEDE. No segundo aspecto, apresenta os problemas técnicos ocasionados com o TEDE, devido à desatualização e instabilidade do *software*, gerando a obsolescência da BDTD/UFG. Diante da situação, deu-se início aos estudos para escolha do *software* a ser utilizado para a migração e configuração dos dados, cujo resultado foi o DSpace, sistema utilizado no Repositório Institucional da UFG (RI/UFG). Foi definida a metodologia utilizada para a correção dos metadados importados do TEDE e a criação de novos *plugins*, para permitir a interoperabilidade dos dois sistemas. As vantagens da preparação para a migração e seus resultados alcançados até o momento são descritos ao final.

Palavras-chave: 1. Biblioteca digital - Migração. 2. Biblioteca digital – UFG. 3. Biblioteca digital - DSpace. 4. Biblioteca digital - TEDE. 5. Interoperabilidade.



ABSTRACT

This work presents two aspects of the Digital Library of Theses and Dissertations (BDTD) of Federal University of Goiás (UFG). The first aspect, describing its creation and the questions imposed on formalizing the guarding process of collections of these material, as well as strategies to solve them. The results demonstrate that BDTD/UFG has more than 2 thousands Electronic Theses and Dissertations (TDEs) in TEDE software, which are of great importance for Graduate Programs and the university community as a whole as well as in promoting Open Access brazilian technical and scientific information. Statistics are also demonstrated during the TEDE's run time. In the second aspect, presents caused technical problems with the TEDE due to outdated software and instability, which led to the inoperability of BDTD/UFG. Given the situation, has begun studies to choose the software to be used for migration and configuration data, which resulted in the DSpace system used in the UFG Institutional Repository (RI/UFG). The methodology used for the correction of imported TEDE metadata and creating new plugins to enable interoperability between the two systems was defined. The advantages of the preparation for migration and the results achieved to date are described at the end.

Keywords: 1. Digital library - Data migration. 2. Digital library - UFG. 3. Digital library - Dspace. 4. Digital library - TEDE. 5. Interoperability.

INTRODUÇÃO

A Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) é um projeto do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) que “integra os sistemas de informação de teses e dissertações existentes nas instituições de ensino e pesquisa brasileiras e estimula o registro de publicação em meio eletrônico” (IBICT, 2008). Em 2005 a Biblioteca Central (BC), unidade do Sistema de Bibliotecas da Universidade Federal de Goiás (SIBI/UFG), e a Pró-Reitoria de Pesquisa e Pós-Graduação¹ (PRPPG), por meio da Câmara de Pesquisa e Pós-Graduação da UFG, iniciaram os estudos para a implantação da Biblioteca Digital de Teses e Dissertações da instituição. A ferramenta utilizada para possibilitar o funcionamento da BDTD foi o Sistema de Publicação Eletrônica de Teses e Dissertações (TEDE), desenvolvido pelo próprio IBICT em duas versões: modular e simplificado.

¹Atualmente desmembrada em Pró-Reitoria de Pós-Graduação (PRPG) e Pró-Reitoria de Pesquisa e Inovação (PRPI).



Desde sua implantação até meados de 2012, o TEDE esteve funcionando normalmente, quando, devido à sua vulnerabilidade ocasionada pela desatualização do mesmo, o sistema foi alvo do ataque de *crackers*², tornando-se obsoleto a partir disso.

Assim, a Gerência de Informação Digital e Inovação (GIDI/BC), responsável pela gestão da BDTD na instituição, iniciou os estudos sobre a escolha do sistema que possibilitasse a migração das teses e dissertações eletrônicas (TDEs) depositadas no TEDE, para outro sistema que utilizasse o protocolo *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) possibilitando a coleta e disseminação dos metadados para repositórios. Primeiramente o módulo Biblioteca Digital, do software SophiA (Prima³) foi estudado para verificação se atendia às necessidades de um gerenciamento de teses e dissertações tal como era feito no TEDE. Após a constatação de que o SophiA não atendia as necessidades, o software DSpace (Duraspace, MIT e HP⁴) foi estudado e definido como satisfatório. Como a UFG possui o Repositório Institucional (RI) e este é gerenciado pelo DSpace, iniciou-se o processo de migração das TDEs que estavam no TEDE para o RI.

Os Repositórios Institucionais são sistemas de informação que possibilitam o acesso à produção da instituição, sem que haja obstáculos de custo para tal, além de oferecem benefícios ao pesquisador, para os administradores acadêmicos, para as instituições e para a comunidade científica (UNIVERSIDADE DE MANCHESTER, 200?, apud LEITE, 2009, p. 23). Eles disseminam a produção intelectual de determinada instituição, possibilitando a disponibilização dos tipos de documentos que forem definidos em sua Política de Informação, como: artigos científicos, trabalhos publicados em eventos, relatórios de pesquisa, teses, dissertações, monografias, dentre outros.

Atualmente a BDTD está em processo de migração para o RI, em fase de finalização e espera-se que até o final de 2014 todas as TDEs sejam disponibilizadas na página do RI e novos registros comecem a ser inseridos pela equipe. Este estudo aborda o processo de preparo e início da migração do banco de dados do TEDE para o RI, as dificuldades encontradas e as alterações realizadas para melhorar o sistema e atender às especificações das teses e dissertações.

² Termo designado a indivíduos que quebram um sistema de segurança de forma ilegal ou antiético.

³ www.prima.com.br

⁴ www.dspace.org



CRIAÇÃO E IMPLANTAÇÃO DA BIBLIOTECA DE TESES E DISSERTAÇÕES DA UFG

Em 2006 o IBICT lançou edital para fomentar a instalação de novas bibliotecas digitais em Instituições Federais de Ensino Superior de todo país. A UFG, dentre várias outras instituições, concorreu e foi selecionada para participar do projeto. A Biblioteca Digital Brasileira de Teses e Dissertações da Universidade Federal de Goiás (BDTD-UFG) foi criada em 29 de junho de 2006 e teve como objetivo disponibilizar via internet as dissertações e teses produzidas pelos programas de pós-graduação da UFG. A versão utilizada do TEDE foi a simplificada.

Para dar início à implantação da BDTD da UFG três unidades acadêmicas foram definidas como piloto. Desta forma os técnicos envolvidos nesta ação teriam maior familiaridade com as tecnologias e os procedimentos necessários à realização da mesma. No piloto foram observados os seguintes problemas: falta de resumos em português e em outra língua; resumos fora das normas; falta de palavras-chave em português e em outra língua; ausência de título em outra língua; ausência da catalogação na fonte ou catalogação fora das normas; formato digital diferente do formato impresso; ausência do termo de autorização; e arquivo protegido.

Além de todas estas questões, houve grande resistência por parte dos autores e de seus orientadores na disponibilização do CPF, necessário para acesso ao TEDE. Observou-se que eram necessárias melhorias na relação entre a biblioteca e os programas de pós-graduação, pois estes não verificavam se os documentos enviados pelos autores estavam condizentes com as normas de publicação da ABNT.

Evolução e estatísticas do TEDE

O objetivo do Sistema de Publicação Eletrônica de Teses e Dissertações (TEDE) é proporcionar a implantação de bibliotecas digitais de teses e dissertações nas instituições de ensino e pesquisa e sua integração a BDTD nacional; gerando bibliotecas digitais institucionais. O *software* é gratuito, distribuído pelo IBICT em um pacote com o Sistema TEDE já implementado, o MTD-BR e a camada do Protocolo OAI-PMH, metodologia de implantação, os manuais operacionais e de usuário, documentação e treinamento. (BDTD, 2006).

A alimentação do TEDE na UFG teve início em 2007 depois dos treinamentos e preparação do material. Até 2009 apenas 81 documentos haviam sido inseridos no sistema, já que a exigência do CPF dos autores ocasionou dificuldades e atraso na alimentação do banco de dados.



Em 2009 a Gerência de Informação Digital e Inovação (GIDI) assumiu a coordenação da BDTD e para otimizar o funcionamento da biblioteca digital, utilizou a seguinte metodologia de trabalho: a Biblioteca Central, em conjunto com a Pró-Reitoria de Pesquisa e Pós-Graduação, trabalhou uma normativa interna descrevendo quais eram as informações essenciais para a disponibilização do documento. Em seguida foram coletadas, por meio de uma tabela separada por programa, todas as pendências referentes a cada dissertação e tese encaminhadas para depósito nas bibliotecas da UFG. De posse deste material foram adotadas duas estratégias para solucionar as pendências: 1- enviar para o coordenador ou secretário de cada programa de pós-graduação, via *e-mail*, a listagem dos trabalhos e respectivos problemas encontrados em cada um juntamente com uma carta explicativa sobre a BDTD; 2- contatar por telefone cada coordenador ou secretário para esclarecer o assunto e solicitar apoio na solução das pendências.

Após um determinado período a equipe envolvida avaliou que as estratégias não apresentaram resultados satisfatórios. Decidiu-se então agendar visitas a cada um dos programas de pós-graduação, o que apresentou resultado muito positivo. Também foi decidido que, para resolver a questão do CPF, o respectivo campo seria retirado do Termo de Autorização e Ciência do autor e a coleta do mesmo seria feita utilizando os dados fornecidos pela própria UFG. Para coletar essa informação, foram traçadas duas estratégias: 1 - usar o Cadastro de Pessoas da UFG; 2 - solicitar, à Pró-Reitoria de Pesquisa e Pós-Graduação da UFG, autorização para consultar o Sistema de Cadastro da Pós-Graduação (SisPG). Como o uso do Cadastro de Pessoas da UFG não foi eficaz, a utilização do SisPG resolveu a questão do CPF.

Devido às pendências encontradas na maioria dos documentos enviados para as bibliotecas da UFG, até março de 2009 constavam apenas 81 documentos depositados na BDTD. Com a retirada do CPF do formulário e a instalação do sistema de consulta da pós-graduação da UFG, em agosto de 2009 notou-se um ganho em quantidade de documentos disponibilizados na BDTD da UFG, assim como nos anos seguintes. O quadro abaixo apresenta o total de documentos inseridos no sistema por ano, de 2008 a 2013.

Quadro 1 - Teses e dissertações inseridas no TEDE de 2008 a 2013

2008	2009	2010	2011	2012	2013	TOTAL
81	417	430	455	574	469*	2326

Fonte: Relatórios GIDI

* Inseridas no TEDE até junho de 2013



Como consequência da desatualização do TEDE⁵ e de problemas técnicos, o banco de dados tornou-se vulnerável e foi alvo da ação de *crackers* em outubro de 2012. O acesso ao público foi removido, ficando somente a inserção de novos documentos. Em setembro de 2013 verificou-se que o TEDE não oferecia diversas funções, tornando-se inoperante. Desta forma, não era possível realizar buscas no TEDE e nem inserir novos documentos.

Insatisfeitos com a situação, estavam os autores, pela não disponibilização dos seus documentos, assim como os servidores que viam parte do seu trabalho acumular, sem poder dar prosseguimento no serviço.

Preparação para migração para o DSpace

Diante da situação, a equipe decide fazer a migração dos dados para outro *software*, sendo necessário, portanto, um estudo que apresentasse qual sistema melhor atenderia às necessidades de uma biblioteca digital de teses e dissertações. Os critérios para escolha do *software* era ser um sistema capaz de submeter, gerenciar, editar e disponibilizar teses e dissertações e possuir o protocolo OAI-PMH⁶ para fazer a coleta e disseminação dos metadados para outros repositórios de dados, permitindo assim, a disseminação dos documentos em âmbito nacional e internacional. Assim, foram analisados os sistemas SophiA (utilizado no Sistema de Bibliotecas da UFG) e o DSpace (utilizado no Repositório Institucional).

Em outubro de 2013 a equipe da GIDI realizou um treinamento *online* sobre o novo Módulo de Biblioteca Digital do sistema SophiA, que poderia ser utilizado para gerenciar as teses e dissertações. A partir desse treinamento, percebeu-se que o sistema não era satisfatório, pois os recursos necessários para alimentação, gestão e disseminação dos dados eram insuficientes à instituição, como o esquema de metadados para teses e dissertações, que descreviam apenas de forma básica esses documentos.

⁵ Os desenvolvedores do TEDE não atualizaram o sistema desde 2010.

⁶ OAI-PMH (*Protocol for Metadata Harvesting*) é um protocolo que permite a interoperabilidade dos conteúdos armazenados em arquivos abertos.



A próxima ação foi verificar se o DSpace atendia tais necessidades e se seria possível migrar o banco do TEDE para esse sistema. O estudo, realizado em fevereiro, demonstrou que o *software* era capaz de gerenciar teses e dissertações, através do processo de *workflow* (submissão, revisão, edição e depósito) de forma prática e segura, além de utilizar opção de embargo e restrição para documentos específicos, como havia no TEDE. O sistema também utiliza o esquema de metadados Dublin Core⁷, que é abrangente e pode ser modificado para melhor atender às necessidades de cada instituição. Além dessas funcionalidades, o DSpace é o *software* do Repositório Institucional da UFG desde 2009, permitindo o depósito dos documentos técnico-científicos institucionais.

Um repositório institucional (RI) “visa ao melhoramento da comunicação científica. Para isso, provêm os mecanismos que aumentam tanto a eficácia da preservação da produção intelectual de pesquisadores e instituições acadêmicos quanto a visibilidade de ambos.” (LEITE; COSTA, 2006, p. 213). Como definição de RI, Lynch afirma: “é um conjunto de serviços que a universidade oferece aos membros de sua comunidade, visando o gerenciamento e disseminação dos materiais digitais criados pela instituição e pelos membros de sua comunidade” (2003, apud LEITE; COSTA, 2006).

Esses serviços têm como característica permitir o armazenamento e a disseminação de tipos de documentos diversificados, tais como: artigos científicos, livros eletrônicos, capítulos de livros, relatórios técnicos, trabalhos publicados em eventos, trabalhos de conclusão de curso, objetos educativos, teses, dissertações e outros.

Entremeio aos estudos a cerca de uma possível migração do TEDE para o DSpace, a equipe é informada que o IBICT estava trabalhando na construção de *plugins* capazes de facilitar e permitir uma migração mais eficaz entre os dois sistemas; o que foi determinante para que a equipe finalmente escolhesse migrar os dados para o repositório institucional.

A partir de então, a equipe se mobilizou no levantamento de estratégias para tornar a migração entre os dois sistemas mais eficaz, assim como facilitar a inserção de novos documentos. Dentre elas:

⁷ É um esquema de metadados para descrição de objetos digitais, tais como, vídeos, sons, imagens, textos e sites na web, composto de 15 elementos básicos, dentre eles (criador, data, assunto, descrição, tipo, formato etc).

- Modificação e especificação dos metadados no repositório institucional, que eram mais básicos do que os existentes no TEDE – novos campos foram adicionados, como: campos para identificação de participantes de banca, orientador e co-orientador, CPF ou matrícula do autor, país, unidades/departamentos, nome do programa de pós-graduação e outros.
- Padronização dos campos de metadados conforme as diretrizes OpenAIRE 2.0⁸, DRIVER 2.0⁹ e do livro ‘Boas práticas para Repositórios Institucionais’¹⁰ (IBICT, 2012) – essas orientações são direcionadas a provedores de conteúdo que visam melhorar a qualidade dos metadados fornecidos durante a catalogação dos objetos digitais, em iniciativas de arquivos abertos. (SHINTAKU; ALMEIDA; MELO, 2011). Nessa etapa, fez-se um estudo acerca da situação do Repositório Institucional em relação às diretrizes acima citadas, que resultou na inserção do campo ‘direitos autorais’ (dc.rights), na adoção da padronização dos títulos dos arquivos conforme o título do documento, correção do campo ‘tipos de documentos’ e no diagnóstico da digitalização para verificação da qualidade dos arquivos digitalizados. Algumas recomendações necessitam ainda serem estudadas e implementadas.
- Identificação dos problemas técnicos de infraestrutura do servidor onde opera o DSpace.

A preparação para o processo permitiu que alguns erros fossem percebidos e solucionados antes da migração, tais como: problema de tradução que havia na versão 3.2 do DSpace, na etapa de preenchimento dos metadados e correção do serviço de *harvesting*¹¹, que não estava funcionando adequadamente.

O IBICT auxiliou no procedimento de customização e migração do TEDE para o DSpace, sendo que essa cooperação foi de extrema importância no sucesso da migração e importação, visto que para que esse ideal ocorra, é necessário conhecimento aprofundado no sistema e um aplicativo que permita a comunicação entre os dois *softwares*.

⁸ <http://www.openaire.eu/>

⁹ <http://www.driver-support.eu/>

¹⁰ <http://livroaberto.ibict.br/handle/1/703>

¹¹ Coleta dos metadados para disseminação.



MATERIAIS E MÉTODOS

A preparação para o processo de migração das teses e dissertações envolveu esforços por parte da equipe da GIDI (SIBI/UFG), do Cercomp (Centro de Recursos Computacionais) e do IBICT. O IBICT desenvolveu e customizou um *plugin* capaz de migrar todos os documentos que estavam no TEDE para o DSpace, versão 4.0. A migração foi realizada em ambiente de testes e depois no servidor do repositório, sendo este armazenado no *data center* da UFG.

Foram realizadas três etapas para que o procedimento obtivesse êxito:

1. Preparação para a migração: nesta etapa foram estudados os dois sistemas e o IBICT desenvolveu e customizou o *plugin* que permitisse a comunicação entre os dois *softwares*. Foram realizados testes em março, que demonstraram sucesso, porém algumas melhorias ainda eram necessárias no aplicativo.
2. Processo de migração: após a finalização do *plugin*, novos testes foram realizados, demonstrando que havia uma maior compatibilidade entre os sistemas e os metadados eram mais específicos para teses e dissertações.
3. Processo de importação e finalização dos documentos: todos os documentos migrados do TEDE encontravam-se na opção de Importação, pelas ferramentas do administrador do repositório (DSpace). Sendo assim, após a finalização da migração, ainda era necessário que as coleções fossem criadas no DSpace, de forma que o Repositório de Teses e Dissertações da UFG tivesse a mesma estrutura que o Repositório Institucional, sendo:

comunidades – regionais das cidades;

subcomunidades – unidades acadêmicas;

subcomunidades – programas de pós-graduação e

coleções – tipos de documentos (mestrado em...
doutorado em...)

Após a criação da estrutura, todos os documentos aglomerados na coleção própria da migração foram importadas e movidos para suas coleções respectivas.



CONSIDERAÇÕES FINAIS

Com a invasão do TEDE confirmou-se sua vulnerabilidade e serviu de alerta para o estudo de um sistema que melhor atendesse a demanda tanto da biblioteca com dos usuários, porque havia muitas críticas em relação a estratégia de busca e inserção de dados no sistema. A definição pelo DSpace foi natural depois que o software gerenciador SophiA não atendeu às expectativas da equipe. Foi gratificante para a UFG ser piloto junto ao IBICT na migração do TEDE para o DSpace. A execução da migração obteve uma taxa de acertos de 96%.

REFERÊNCIAS

BUENO, Cláudia Oliveira de Moura; SILVA, Odete Jacomini da; FERREIRA, Liliane Juvência Azevedo. Biblioteca digital de teses e dissertações e os programas de pós-graduação: uma interação necessária. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS - SNBU, 16., 2010, Rio de Janeiro. *Anais eletrônicos* do XVI Seminário Nacional de Bibliotecas Universitárias, Rio de Janeiro: UFRJ, 2010a. n.p. Disponível em: <<http://www.bc.ufg.br/sophia/bc/publicacoes/bdtd.pdf>>. Acesso em: 16 jul. 2014.

_____ et al. O Portal de Periódicos da UFG. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS - SNBU, 16., 2010, Rio de Janeiro. *Anais eletrônicos* do XVI Seminário Nacional de Bibliotecas Universitárias, Rio de Janeiro: UFRJ, 2010b. n.p. Disponível em: <http://www.gapcongressos.com.br/eventos/z0070/trabalhos_pesquisa.asp?pag=1>. Acesso em: 20 jun. 2014.

CUNHA, Murilo Bastos. Construindo o futuro: a biblioteca universitária brasileira em 2010. *Ciência da Informação*, Brasília, v. 29, n.1, p.71-89, jan./abr. 2000.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Biblioteca Digital Brasileira de Teses e Dissertações. Brasília: IBICT, 2010. Disponível em: <<http://www.ibict.br>>. Acesso em: 25 jun. 2014.

KURAMOTO, Hélio. *BDTD: uma questão de interoperabilidade humana?* [Online]. Disponível em: <<http://kuramoto.blog.br/2006/04/14/bdtd-uma-questao-de-interoperabilidade-humana>>. Acesso em: 20 ago. 2014.

LEITE, Fernando César Lima. *Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto*. Brasília, DF: IBICT, 2009. Disponível em: <http://www.ibict.br/anexos_noticias/repositorios.institucionais.F.Leite_atualizado.pdf>. Acesso em: 15 jul. 2014.



_____; COSTA, Sely. Repositórios institucionais como ferramentas de gestão do conhecimento científico no ambiente acadêmico. *Perspect. ciênc. inf.*, Belo Horizonte, v.11 n.2, p. 206 -219, mai./ago. 2006.

MARCONDES, Carlos H.; SAYÃO, Luis F. Interação e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: proposta da Biblioteca Digital Brasileira. *Ciência da Informação*, Brasília, v. 30, n.3, p. 24-33, set./dez. 2001.

_____; _____. À guisa de introdução: repositórios institucionais e livre acesso. In: SAYÃO, L. F. et al. *Implantação e gestão de repositórios institucionais : políticas, memória, livre acesso e preservação*. Salvador: EDUFBA, 2009. p. 9-19.

MEADOWS, A.J. *A comunicação científica*. Brasília: Briquet de Lemos/Livros, 1999.

SILVA, Odete Jacomini da. *Projeto para implantação da biblioteca digital de teses e dissertações na Universidade Federal de Goiás – TEDE-UFG*. Goiânia: UFG, 2006.



INTERFACES MÓVILES EN DSPACE*

Sergio Nieto-Caramés¹ y Emilio Lorenzo-Gil²

¹ Ingeniero Informático responsable de Desarrollos Dspace en la empresa Arvo Consultores y Tecnología SL Polígono Industrial de Olloniego, parcela B51, 33660 Oviedo, Asturias, España. Tel. +34 985 98 29 03. snieto@arvo.es.

² Ingeniero Telecomunicaciones director de la empresa Arvo Consultores y Tecnología SL Polígono Industrial de Olloniego, parcela B51, 33660 Oviedo, Asturias, España. Tel. +34 985 98 29 03. elorenzo@arvo.es.

RESUMEN

La utilización de dispositivos móviles para el acceso a internet es cada vez más frecuente gracias a la reducción de tarifas y a la disponibilidad de aparatos cada vez más funcionales. Este incremento de uso implica replantearse los servicios ofrecidos por las Bibliotecas Universitarias. El desarrollo de interfaces móviles para las Bibliotecas empieza a ser importante, estimándose que un 44% de las bibliotecas universitarias ofrecen algún tipo de servicio servicios móviles a sus usuarios. Las características de estos dispositivos, principalmente tamaño de pantalla e interacción de usuario, requiere la adaptación de las interfaces web para facilitar su interacción con el usuario en estos nuevos dispositivos. En la comunicación explicaremos las distintas alternativas que el software Dspace cuenta para poder mejorar la interacción y experiencia de uso con estos dispositivos. Introduciremos las técnicas existentes para adaptar las páginas y contenidos a los dispositivos móviles. Igualmente, revisaremos los mecanismos básicos y claves de adaptación que usa DSpace para trabajar con los aspectos móviles desde la versión 1.8 en adelante y señalaremos los mecanismos que se pueden usar en Dspace para detectar el tipo de dispositivo y la redirección a múltiples versiones de las páginas. Seguidamente revisaremos la nueva interfaz adaptativa de Dspace, basada en herramientas de diseño responsivo, como Bootstrap, incluidas en el tema Mirage2 y las características funcionales logradas con esta interfaz. Finalmente mostraremos las pautas de desarrollo de aplicaciones para dispositivos Android que interactúan con Dspace en funcionalidades específicas, como la realización de depósitos mediante protocolo SWORD.



ABSTRACT

The use of mobile devices for accessing the Internet is becoming more frequent every day, due to reduced prices and the availability of increasingly functional devices. This increased usage requires to evaluate the services offered by University Libraries. The development of mobile interfaces for Libraries is becoming important, being estimated that a 44% of university libraries offer some kind of mobile service to its users. The characteristics of these devices, mainly screen size and user interaction, requires the adaptation of the web interface to facilitate its interaction with the user in these new devices. We will show the possibilities that DSpace software provides in order to be able to improve the interaction and experience of use with these devices. We will introduce new ways to adapt websites and contents to a mobile version, together we will review the basic mechanisms that DSpace uses to work with the mobile aspects from version 1.8 forward, offering an structural vision of the mobile subject in DSpace, the mechanisms used to detect the right device type and the redirection to the suitable version of the webpage. We will revise the new adaptive interface of DSpace, based on responsive design tools, as Bootstrap, included in the Mirage 2 theme, and the functional characteristics accomplished with this new interface. Finally, we will show how to develop applications for Android mobile devices interacting with DSpace in specific functionalities, as the accomplishment of deposits using SWORD protocol.

Palabras clave/ Keywords: Repositorios Institucionales, Institutional Repositories, Mobile Interface Interface Móvil, Smartphone, Dspace

INTRODUCCIÓN

Los servicios móviles son ya parte esencial de nuestra experiencia como receptores de información y usuarios de la web y los dispositivos móviles avanzados, smart-phones, constituyen la forma preferida de acceso de las jóvenes generaciones.

Las bibliotecas universitarias, han sido pioneras en la adopción de interface móviles en las Universidades, centrándose inicialmente en servicios del tipo catálogo, información de servicios, sistemas de soporte a usuarios, proporcionando de forma efectiva soluciones y servicios a los estudiantes y resto de comunidad académica. Dada la variedad de aplicaciones desplegadas, puede afirmarse que las bibliotecas universitarias están en un nivel suficiente de madurez y sostenibilidad de las soluciones móviles.



En este contexto, en los últimos años, han eclosionado una variedad de alternativas para el desarrollo de interfaces móviles para repositorios DSpace. Así, se solicita por la comunidad de usuarios en la lista de funcionalidades requeridas en septiembre de 2011, aparece en verano de 2012 el primer tema móvil, y éste se incorpora como solución beta a DSpace v3, finales de 2012 y como funcionalidad soportada a DSpace v4, finales de 2013. En paralelo, DSpace-JSPUI v4 incorpora un interface adaptativo de forma nativa y la empresa @atmire anuncia en la conferencia Open Repositories en junio de 2014 la disponibilidad de su tema Mirage2 para Dspace-XMLUI v5 de forma nativa y con posibilidad de incorporarlo a las versiones DSpace v3 y v4.

Ante esta variedad de alternativas, el dilema principal en la provisión de servicios móviles en DSpace es decidir si optar en crear interfaces específicos para los diferentes segmentos de dispositivos móviles en el mercado o bien crear un sitio web capaz de interactuar con dispositivos tradicionales a la par que ofrecer contenidos a los usuarios con dispositivos móviles.

Compararemos en primer lugar las tres opciones principales que se presentan en la actualidad para el desarrollo de aplicaciones móviles:

- a) Interface específica, URL dedicada y exclusiva a móviles, con formato del tipo m.repositorio.com y diseño independiente.
- b) Diseño responsivo o adaptable, conocido por las siglas RWD (del inglés, Responsive Web Design), que permite que con un único diseño y una única URL tengamos una visualización adecuada en cualquier dispositivo, adaptando la apariencia de las páginas web al dispositivo que se esté utilizando para visualizarla.
- c) Entrega dinámica o **Dynamic Serving es una** evolución del concepto de diseño responsivo que permite optimizar el interface para adaptarlo a un rango mas amplio de dispositivos.

A continuación revisaremos las soluciones específicas que para los anteriores opciones existen en DSpace y por último exploraremos las posibilidades de construir interfaces específicos, mediante apps, aplicaciones móviles, adaptados a funcionalidades concretas de una instalación DSpace.

ALTERNATIVAS DE DESARROLLO DE INTERFACES MÓVILES

Interface específica Web dedicada y exclusiva a móviles. Tienen normalmente el formato de URL m.repositorio.com. El código HTML generado para esta interface difiere del correspondiente a la interface principal.

Diseño responsivo o adaptable o adaptativo, conocido por las siglas RWD (del inglés, Responsive Web Design), que permite con un sólo diseño web tengamos una visualización adecuada a cualquier dispositivo. Su filosofía de diseño y desarrollo tiene como objetivo adaptar la apariencia de las páginas web al dispositivo que se esté utilizando para visualizarla.

Entrega dinámica, es una evolución del concepto de diseño responsivo, permite optimizar aún más el interface, ya que permite seleccionar qué elementos se requiere presentar en pantalla en función de parámetros como tamaño de pantalla o tipo de dispositivo.

	Interface específica móvil	Diseño Adaptativo	Entrega Dinámica
Ventajas	<p>Carga y navegación más rápida</p> <p>Posibilidad de personalizar considerablemente la interface por usar un contexto distinto.</p> <p>Contenido mucho mas adaptable a cada tipo de usuario / dispositivo.</p>	<p>URL e interface únicos, válidos para múltiples dispositivos (PC, móvil y tabletas).</p> <p>Una vez desarrollado, es simple la optimización para nuevos dispositivos.</p> <p>Simplificación del SEO, pues la web es única. Recomendado explícitamente por Google (criterios SEO)</p>	<p>Mejora las posibilidades de personalización de del modelo adaptativo.</p>
Inconvenientes	<p>Duplica el trabajo, pues son dos interfaces distintas con diferente dominio.</p> <p>Doble configuración de los ficheros SEO: robots.txt, sitemap.xml, etc..</p> <p>Necesidad de implementar etiquetas switchboard para evitar la penalización SEO por duplicación de contenidos</p>	<p>Dificultad de adaptar o evolucionar a esta tecnología a los interfaces (Look&feel) ya existentes.</p> <p>Curva de aprendizaje compleja, uso de nuevas técnicas CSS: CSS3, SASS, SCSS.</p>	<p>Proceso complejo de selección de opciones de contenidos dinámicos</p> <p>Curva de aprendizaje aún mas compleja: CSS authoring frameworks, nuevas metodologías de diseño de interfaces, nuevas técnicas SEO</p>

Tabla 1: Comparativa de alternativas de desarrollo de interfaces móviles

INTERFAZ DSPACE-XMLUI ESPECÍFICA PARA DISPOSITIVOS MÓVILES

“Los usuarios móviles tienen típicamente intereses diferentes a los de los dispositivos PC. Probablemente tienen intenciones más directas e inmediatas que los usuarios tradicionales de la web. Habitualmente sus intereses comprenden encontrar elementos específicos de información relevantes a su contexto” (W3C, 2008).

El primer tema móvil disponible en el código de DSpace aparece en 2012 y se enfoca a los usuarios finales del repositorio, asumiendo, en línea con lo planteado unos años antes por el W3C, que sus tareas principales son de búsqueda, navegación y lectura de contenidos.

Siguiendo el criterio de que los sitios web diseñados específicamente para dispositivos móviles mejoran su usabilidad, el enfoque en ese momento fue crear un interfaz específico para DSpace. Igualmente el mercado 4G era inexistente y los dispositivos móviles avanzados no eran aún predominantes en ese momento.

El tema móvil, disponible para DSpace versión3, hace uso de la característica de XMLUI de poder usar diferentes temas seleccionables según contexto. Un tema de XMLUI es una capa intermedia de código que permite a los desarrolladores crear una interface personalizada para un sitio DSpace, una comunidad o una colección.

Esta funcionalidad de selección de tema, que como mencionamos es particular de XMLUI, permite diseñar, desarrollar y desplegar el interface móvil como el resto de temas disponibles, pues tiene la misma estructura de carpetas y ficheros

```
+-- mobile
| +-- lib
| | +-- cookies.js
| | +-- detectmobile.js
| | +-- images
| | | +-- ajax-loader.gif
| | | +-- default-thumbnail.png
| | | +-- icons-18-black.png
| | | +-- icons-18-white.png
| | | +-- icons-36-black.png
| | | +-- icons-36-white.png
| | +-- m-tweaks.css
| | +-- sc-mobile.css
| | +-- sc-mobile.min.css
| | +-- mobile.xsl
| | +-- sitemap.xmap
| | +-- themes.xmap
| +-- readme.txt
```

Figura 1: Estructura del tema Mobile, Dspace v3

Como característica reseñable del tema, incorpora un javascript (detectmobile.js), de detección del tipo de dispositivo (móvil o navegador PC) desde el que se realiza el acceso a DSpace, aplicando el interfaz-tema mobile en vez del genérico del sitio web.

Para realizar las pruebas y ajustes del diseño de este interfaz no se requiere el uso específico de un dominio móvil, ni de un dispositivo móvil, ya que se puede activar mediante la reconfiguración de nuestro fichero xmlui.xconf, determinante de los temas usados por el interfaz XMLUI de DSpace.

El inconveniente principal de esta aproximación a los proyectos móviles es la dificultad de mantener la compatibilidad y el soporte a los múltiples tipos de dispositivos, sistemas operativos con su soporte variable a los estándares HTML y CSS , tamaños de pantalla diversos, etc. en un entorno de auténtica explosión del mercado móvil. El soporte completo para el mercado de tablets, phablets, smartphones 4G, *smartphones 3G* básicos, terminales con funciones avanzadas, con mas de una decena de sistemas operativos compitiendo y evolucionando, parece simplemente inviable, siendo poco práctico crear una nueva versión del interfaz para cada combinación o variante principal.

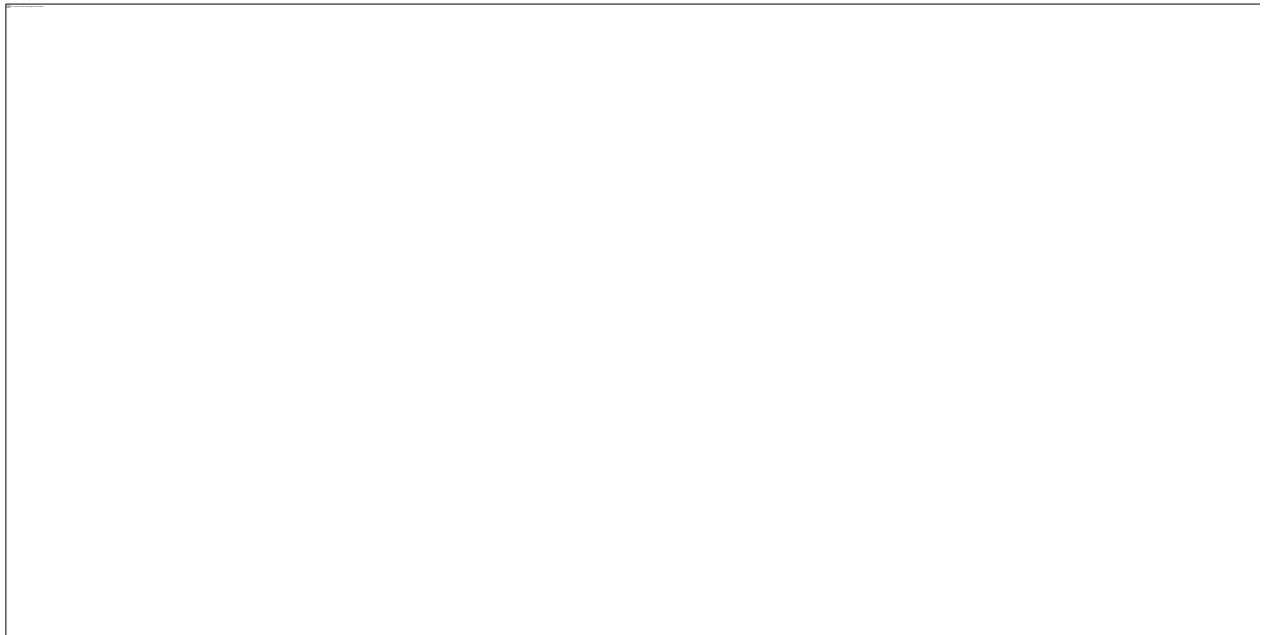


Figura 2: DSpace v3 con tema específico móvil en emuladores iOS, Android, Opera y Windows Phone. Figura de (Tzoc, 2013)



DISEÑO RESPONSIVO EN DSPACE XMLUI, EL TEMA MIRAGE2

La respuesta al panorama que se apuntaba en el párrafo anterior viene de la mano de los conceptos de Diseño Web Adaptativo, que básicamente es un enfoque al diseño web en el que éste se orienta a independizar la experiencia del usuario del dispositivo que éste usa.

En los dos últimos años, la comunidad DSpace ha estado explorando este paradigma de diseño web y evaluando la viabilidad de un diseño web responsivo para la interface de usuario. Los resultados parecen ser positivos, pues a finales de 2003, DSpace-JSPUI v4 incorpora un interface adaptativo de forma nativa, discontinuando el aspecto “clásico” de las versiones anteriores del interfaz JSPUI. Igualmente, en junio de 2014, aunque precedido por algunas noticias previas, la empresa @atmire anuncia en la conferencia Open Repositories 2014, la disponibilidad del tema responsivo denominado Mirage2, para el interfaz XMLUI que puede usarse en las versiones v3 y v4, y que será incluido en la distribución de código de la versión 5, prevista para finales de 2014.

El interfaz Mirage2, basándose en los principios de diseño adaptativo o responsivo, los extiende y evoluciona hacia un modelo de Entrega Dinámica, proporcionando un entorno responsivo, adaptativo y fluido, para lograr un entorno a la vez común y específico a cada dispositivo que usemos para acceder al repositorio.

Mirage 2 no es solo un tema adicional a incorporar a la carpeta de temas seleccionables en el diseño inicial del repositorio, como pueden ser los temas Classic, Manankin, Kubrick o Mirage, sino que se apoya y requiere una serie de herramientas facilitadoras y nuevos estándares de desarrollo web, como son:

- SASS, lenguaje de maquetación basado en una extensión de CSS3 que nos permite definir variables, importar propiedades CSS, usar operadores y otras utilidades.
- Bootstrap 3, marco de trabajo, framework, para desarrollar interfaces responsivos basados en lenguajes orientados a Web como HTML, CSS y Javascript. Su ventaja principal es la generación de diseños responsivos de manera fácil e intuitiva a partir del uso de rejillas de diseño (grids) y control responsivo de la rejilla.
- Compass, *CSS Authoring Framework que de forma similar a Bootstrap* nos permite desarrollar interfaces responsivos para nuestro sitio Web.

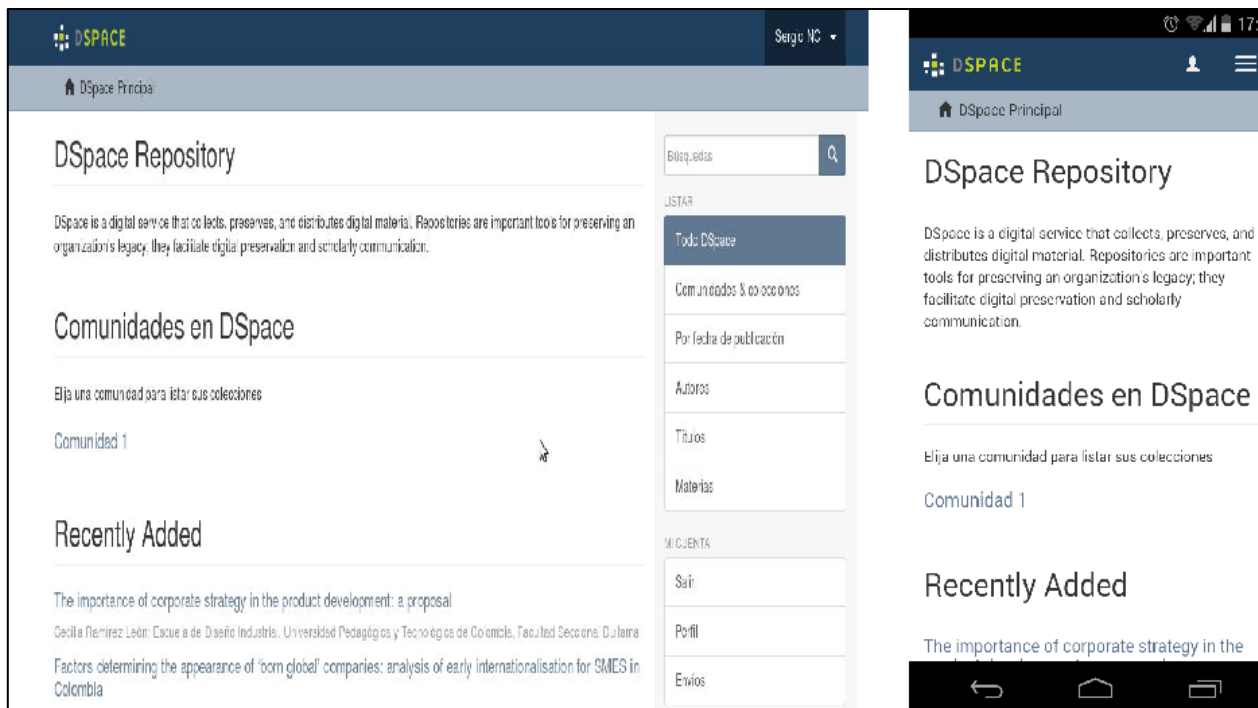


Figura 3: DSpace v4 con tema Mirage2 en navegador Chrome en PC y en un emulador de Android.

DESARROLLO DE APLICACIONES MÓVILES ESPECÍFICAS

Aunque un sitio web adaptado a móvil sea funcionalmente bueno, una aplicación móvil puede ser aún mejor. Se han medido tasas de satisfacción del 76% en el uso de aplicaciones móviles, significativamente mayor que el 64% registrado para webs específicamente adaptadas a móviles. En especial, se recomienda la creación de aplicaciones móviles cuando existan servicios adaptados u orientados al uso móvil. (Nielsen, 2011)

Aunque no se ha requerido por la comunidad DSpace la necesidad de aplicaciones móviles cubriendo funcionalidades específicas, en la frontera del uso de los repositorios DSpace siempre han existido aplicaciones web orientadas a cubrir requisitos funcionales limitados, en contraposición a la funcionalidad amplia del interface general del aplicativo.

Estas aplicaciones han tenido un foco, aunque no exclusivo, en el proceso de depósito de objetos al repositorio, usando las capacidades de integración que DSpace tiene a través de sus interfaces REST (Representational State Transfer) o de forma predominante, SWORD (Simple Web-service Offering Repository Deposit).



Para el objeto que nos ocupa, explorar la viabilidad de aplicaciones móviles para DSpace, se diseñó una app móvil para dispositivos Android, limitada a la funcionalidad de depósito simple en el repositorio.

Los criterios de diseño fueron simplicidad de uso, orientación a la tarea de depósito simplificado de objetos en DSpace y uso de los estándares técnicos, implementando a través de la API de Google las librerías JAVA de soporte al protocolo SWORD.

Teniendo en mente este criterio de simplicidad y orientación a la tarea, la aplicación consta básicamente de los siguientes elementos:

- Pantalla inicial de configuración, que implementa los procesos de identificación de la plataforma DSpace de conexión, autenticación del usuario depositante y negociación del service-document, que recordemos es la primera respuesta del protocolo SWORD ante una solicitud de depósito.
- Formulario de entrada de datos, que recoge unos metadatos descriptivos mínimos (simplicidad), permite la subida de uno o varios ficheros y paquetiza esos elementos en un fichero METS en formato comprimido, que será aceptado por el interface SWORD v1 del DSpace destino.

REFERENCIAS BIBLIOGRÁFICAS

Tzoc, E. (2013). A Mobile Interface for DSpace. In *D-Lib Magazine March/April 2013 Volume 19, Number 3/4* doi:10.1045/march2013-tzoc

World Wide Consortium, W3C (2008). Mobile Web Best Practices 1.0, Basic Guidelines. W3C Recommendation
<http://www.w3.org/TR/mobile-bp/> Accessed July 30, 2014

Nielsen, J. (2011). Mobile Usability Update.
<http://www.nngroup.com/articles/mobile-usability-update/> Accessed July 30, 2014

REPOSITARIOS DSPACE CON MÚLTIPLES CONTEXTOS OAI-PMH

Marisa Raquel De Giusti¹, Facundo Gabriel Adorno², Ariel Jorge Lira³

¹Ingeniera en Telecomunicaciones y Profesora de Letras. Directora del Servicio de Difusión de la Creación Intelectual (SEDICI) y del Proyecto de Enlace de Bibliotecas (PREBI), Universidad Nacional de La Plata (UNLP) e Investigador de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICBA).

²Estudiante avanzado en la Facultad de Informática, UNLP. SEDICI, UNLP.

³Licenciado en Informática. SEDICI, PREBI, UNLP.

INTRODUCCIÓN

En la actualidad, casi la mitad de los repositorios digitales del mundo están soportados por el software DSpace[13] y lo mismo sucede en igual o mayor medida con los nuevos repositorios que se siguen creando día a día. Esto puede atribuirse entre otros motivos a que es un desarrollo de código abierto, a la funcionalidad que provee, a la gran cantidad de documentación y experiencias en línea sobre su uso e instalación y en particular a su gran comunidad de usuarios y desarrolladores que constantemente lo actualiza y expande.

El proyecto DSpace posee una arquitectura[12] compleja que se divide en 3 grandes capas o áreas: aplicación, lógica de negocios y almacenamiento. Ver Figura 1.



FIGURA 1 - Arquitectura DSpace

La capa de Aplicación incluye todas las herramientas que permiten al exterior (usuarios u otros sistemas) hacer uso del repositorio; por ejemplo XMLUI, JSPUI, módulo OAI, SWORD Server, entre otros. La capa intermedia mantiene la lógica transversal a todas las aplicaciones y que rige el funcionamiento interno del repositorio. Finalmente la capa de almacenamiento se encarga de todas las tareas específicas de guardado y recuperación desde almacenamiento secundario, es decir, bases de datos y sistema de archivos.



Una instalación DSpace ofrece varios puntos de configuración y extensión mediante diversos archivos, que definen parte del modelo de datos del repositorio, los flujos de trabajo, las características de interfaz y muchísimos puntos más. En ciertos casos, la forma y sintaxis de estos archivos son simples y permiten entender su lógica sin mayor complejidad. Sin embargo, muchos de estos archivos requieren no solo comprender el funcionamiento de la aplicación sino también entender su implementación y otros conceptos vinculados a las herramientas complementarias que utiliza DSpace para funcionar. Algunos ejemplos de éstos casos son la customización del proceso de ingesta o *submission*, del formulario de carga, de la configuración avanzada del módulo de búsqueda, entre otros.

Este trabajo se centra en aquellas tecnologías que brindan un servicio de interoperabilidad sobre el protocolo OAI-PMH[8] mediante el módulo OAI2.0 provisto por DSpace[1]. Dado que la interoperabilidad determina la capacidad de intercambiar información entre dos o más sistemas, este módulo se localiza en la capa de aplicación de DSpace, en contacto con el mundo exterior. De acuerdo a estudios realizados, OAI-PMH es el protocolo más utilizado para implementar interoperabilidad de metadatos en el ámbito de repositorios digitales [15].

OAI-PMH provee un marco de interoperabilidad independiente de la aplicación subyacente y basado en la cosecha de metadatos. El protocolo define 2 roles claros, uno de *data provider*, correspondiente a la fuente que expone metadatos usando el protocolo, y otro denominado *service provider*, que se aplica a los sistemas externos que toman estos metadatos de uno o más *data providers*, los procesan y generan un servicio de valor agregado a partir de los mismos. En ocasiones, un repositorio que posee una red de repositorios colaboradores puede cumplir ambos roles. El módulo DSpace que se responsabiliza de la interoperabilidad mediante OAI-PMH es denominado OAI 2.0.

La experimentación que antecedió la realización de este trabajo se realizó en el repositorio institucional de la UNLP: el Servicio de Difusión de la Creación Intelectual (SEDICI)[3]. Actualmente, el repositorio cuenta con gran cantidad de publicaciones, más de 35 mil obras, que se exponen desde el módulo OAI 2.0 bajo 3 contextos de cosecha, correspondientes a las directrices de interoperabilidad OpenAIRE[4][5], DRIVER[6] y SNRD[7]. En las secciones siguientes se incluye un breve análisis sobre las decisiones tomadas y configuraciones usadas para la adecuación de este módulo a las 3 directrices de interoperabilidad.

DIRECTRICES

Las directrices analizadas están muy vinculadas dado que todas comparten a DRIVER2 como raíz y por tanto presentan similitudes en sus requerimientos y estructura. Ver Figura 2.

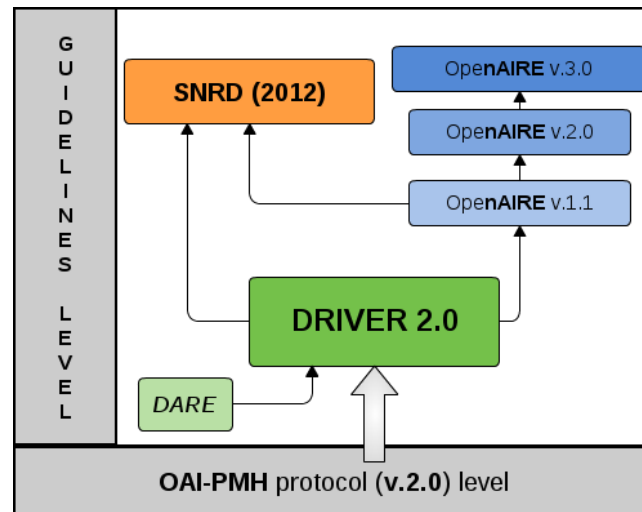


FIGURA 2 - Relaciones entre directrices

De este modo, varios de los aspectos de algunas de estas directrices están basados en los aspectos de otras; por ejemplo, las tipologías documentales usadas en DRIVER son también utilizadas por OpenAIRE. Según se indica en sus especificaciones, las directrices SNRD toman elementos de las directrices DRIVER y OpenAIRE 1.1; y las directrices OpenAIRE 1.1 se basan en DRIVER —las actuales directrices OpenAIRE 3.0 también—. Consecuentemente, SNRD y OpenAIRE 1.1 presentan características muy similares y se asemejan mucho a DRIVER.

Un objetivo común en estas directrices es la normalización de la representación de algunos metadatos y el cumplimiento de ciertos metadatos de forma obligatoria, recomendada u opcional. Para ello, en sus respectivas especificaciones se define un formato de metadatos común de interoperabilidad, Dublin Core, sobre el cual se deben plasmar los metadatos de acuerdo a reglas sintácticas de obligatoriedad, dominio de datos y orden de metadatos. Para refinar aún más el dominio de los datos expuestos, se determinan y utilizan vocabularios específicos (p.e. info:eu-repo/semantics/) o códigos estándares (p.e. ISO 639-3). Esta estandarización en la representación y codificación de metadatos, además de los protocolos de comunicación, permite la interoperabilidad entre los distintos sistemas de información[16].

En la TABLA 1 se enuncian algunas de las diferencias y similitudes entre los aspectos de las directrices analizadas.

	DRIVER 2.0	SNRD 2012	OpenAIRE 1.1	OpenAire 3.0
Definición de set OAI-PMH:	<i>setName</i> : Open Access DRIVER set" <i>setSpec</i> : driver	<i>setName</i> : Sistema Nacional de Repositorios Digitales <i>setSpec</i> : snrd	<i>setName</i> : EC_fundedresources set <i>setSpec</i> : ec_fundedresources	<i>setName</i> : OpenAIRE <i>setSpec</i> : openaire
Condiciones de acceso al contenido	(a) en acceso abierto (b) con embargo Se recomienda no exponer contenido embargado.	(a) acceso abierto (b) con embargo.	(a) en acceso abierto (b) con embargo, (c) restringidos	(a) en cceso abierto, o (b) resultantes de algún proyecto de investigación financiado, independientemente de su estado de acceso
Uso de <dc:rights>	<i>Uso</i> : Recomendado. <i>Esquema</i> : no especificado	<i>Uso</i> : Obligatorio. <i>Esquema</i> : vocabulario OpenAIRE 1.1 para niveles de acceso. Se debe exponer como la primera instancia.	<i>Uso</i> : Recomendado. <i>Esquema</i> : vocabulario propio de 4 niveles de acceso: closedAccess, embargoedAccess, restrictedAccess, openAccess.	Ídem OpenAIRE 1.1.
Uso de <dc:format >	<i>Uso</i> : Recomendado <i>Esquema</i> : MIME de IANA.	<i>Uso</i> : Obligatorio <i>Esquema</i> : MIME de IANA.	Ídem DRIVER.	Ídem DRIVER.
Tipologías y versiones - <dc:type>	<i>Uso</i> : 1. Tipo Driver (Obligatorio) 2. Otros tipos (Opcional) 3. Versión de la obra (Opcional) <i>Esquema</i> : define 2 vocabularios propios para tipos y versiones en info:eu-repo/semantics.	<i>Uso</i> : 1. Tipo Driver (Obligatorio) 2. Tipo SNRD (Obligatorio) 3. Versión de la obra (Opcional) <i>Esquema</i> : usa la misma tipología que Driver y define una propia para tipos y obligatoriedad: (1) - TIPO Driver (Obligatorio) (2) - TIPO SNRD (Oblig.) (3) - Versión TIPO Driver (Opcional)	Ídem a DRIVER.	Ídem a DRIVER.

TABLA 1 - Algunas diferencias y similitudes entre las directrices analizadas.



Resulta conveniente considerar todas éstas cuestiones comunes entre las directrices, para poder determinar qué particularidades se puede reutilizar en las implementaciones de cada contexto.

Una extraña diferencia surge en relación a la inclusión de la fecha de fin embargo. Por un lado SNRD recomienda indicar dicha fecha en la segunda instancia del metadato dc.rights, ej:

```
<dc:rights>2016-12-31</dc:rights>
```

Por otro lado, OpenAire recomienda utilizar una instancia del metadato dc.date bajo la sintaxis 'info:eu-repo/date/embargoEnd/AAAA-MM-DD', Ej:

```
<dc:date>info:eu-repo/date/embargoEnd/2016-12-31</dc:date>
```

MÓDULO OAI 2.0

El módulo OAI2.0[1] está desarrollado en base a la librería XOAI Java Toolkit[2] y pensado específicamente para ser usado en repositorios DSpace[9] como data provider OAI-PMH.

Esta aplicación presenta una configuración flexible que permite su adaptación a diversas necesidades, ya sea para cumplir con políticas institucionales, con directrices de interoperabilidad u otros criterios arbitrarios. Su implementación se basa en 4 conceptos fundamentales: Contexto, Filtro, Transformador y Mapeador.

Un *contexto* OAI 2.0 denota una selección arbitraria de registros del repositorio que cumplen con un conjunto de características determinadas. Existe un contexto por defecto denominado *request* que incluye todos los recursos del repositorio.

Los *filtros* determinan qué ítems del repositorio deben ser incluidos en el contexto solicitado. En su implementación, cada filtro sabe 2 cosas: I) qué características deben tener los registros a incluir en el contexto y II) si dado un registro, este debe ser incluido o no en el contexto. Entonces, cada filtro:

- I. determina las características de los recursos a preseleccionar a partir de una o más condiciones sobre los metadatos. Esta condición se deberá generar tanto para código de consulta SQL como de consulta Solr. Luego, estas condiciones se utilizarán en combinación con todas las cláusulas de todos los filtros para determinar el conjunto base de recursos que se usará para los siguientes pasos.
- II. determina si un ítem es válido en el contexto a partir del análisis de los datos contenido en el ítem, ya sean metadatos descriptivos, administrativos o incluso sus archivos.

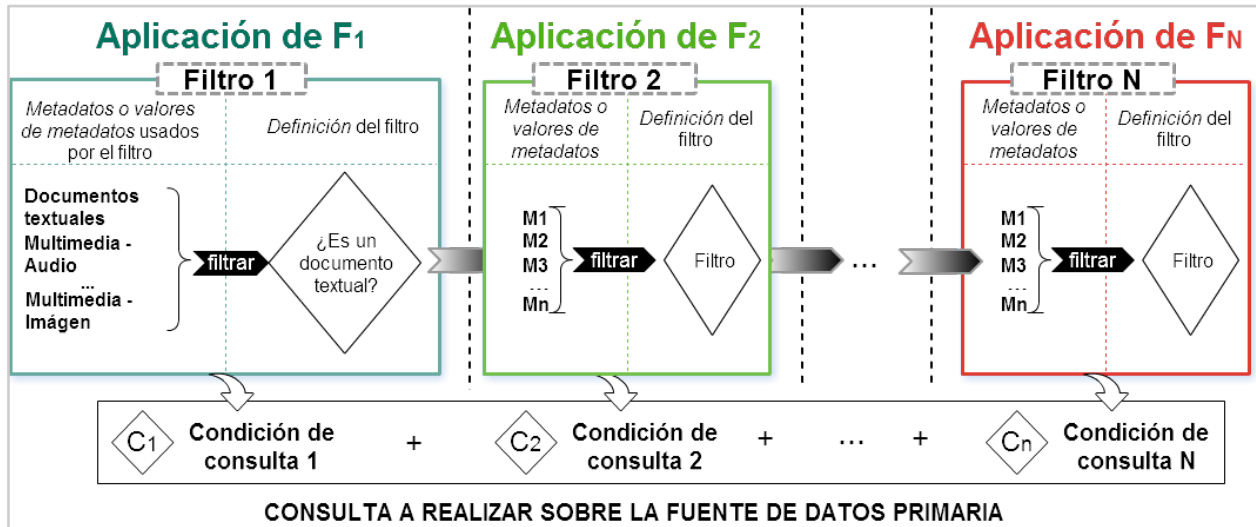


FIGURA 3 - Diagrama explicativo sobre el funcionamiento de los filtros en OAI 2.0

Un *transformador* actúa directamente sobre los metadatos de los documentos filtrados, convirtiendo un conjunto de metadatos de entrada en otro conjunto de metadatos de salida. Una transformación puede implicar tanto una reducción como incremento en los datos de salida, así como también una alteración de valores de dominio.

El *mapeador* toma la salida del transformer y la modifica de acuerdo al formato de metadatos requerido, típicamente oai_dc.

Como se muestra en la TABLA 2, la utilización de estos componentes varía según el verbo OAI-PMH en uso.

¿Qué componentes se utilizan para responder cada uno de las solicitudes OAI-PMH?	Filtro	Transformador	Formato
Identify	No	No	No
ListRecords	Sí	Sí	Sí
GetRecord	Sí	Sí	Sí
ListSets	No	No	No
ListMetadataFormats	No	No	No
ListIdentifiers	Sí	No	No

TABLA 2 - Relación entre verbos OAI-PMH y componentes XOAI

CONFIGURACIÓN

DSpace permite personalizar el funcionamiento de OAI 2.0 mediante varios archivos ubicados en la sección de configuración, es decir, en el directorio *config* de la instalación. En la figura 4 se puede ver la estructura de archivos de datos y principales archivos para configurar y personalizar el módulo.

El archivo de propiedades *oai.cfg* especifica las propiedades elementales de configuración del módulo, como ser el tipo de almacenamiento desde la que se procesarán los metadatos de los ítems del repositorio; puede usarse la base de datos o Solr. En este último caso, debe completarse también la propiedad *solr.url* con la URL de acceso al índice oai dentro de Solr.

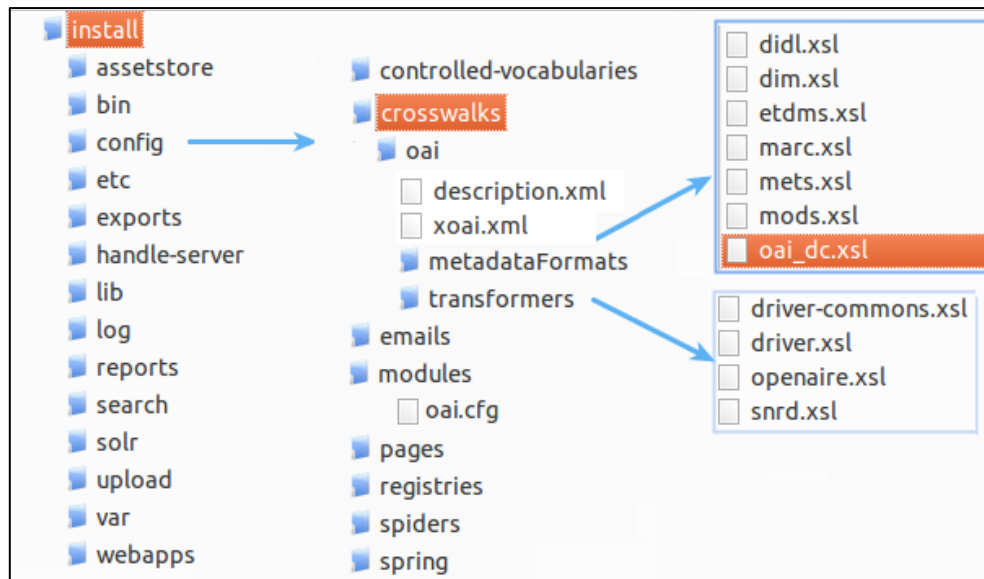


FIGURA 4 - Estructura de directorios DSpace 4 y archivos de configuración OAI 2.0

El archivo de configuración *xoai.xml* es el punto principal para armar y enlazar todos los componentes del módulo, definiendo los contextos habilitados, filtros, transformadores y formatos de metadatos disponibles en cada contexto.

El directorio *transformers* contiene hojas de estilo de transformación definidas en lenguaje XSLT[10], que suelen usarse para agregar y eliminar metadatos, para alterar su orden o realizar cualquier otra modificación sobre los mismos.

Finalmente, el directorio *metadataTransformers* reúne las hojas de estilo XSLT para mapeo de metadatos a un formato de metadatos específico. Estos archivos son también referenciados desde *xoai.xml*. A diferencia de los *transformers*, su función es transformar los el formato interno de metadatos original a un formato de metadatos en particular. El formato deseado se determina en función del parámetro *metadataPrefix* del de la solicitud OAI-PMH, y de la configuración definida en *XOAI.xml* para el contexto seleccionado.

FUNCIONAMIENTO EN ETAPAS

El procesamiento de una solicitud OAI-PMH en el módulo OAI 2.0 implica el procesamiento de los componentes enunciados en el inciso anterior. Los componentes puntuales invocados en cada solicitud dependen exclusivamente del contexto seleccionado y de su configuración.

A continuación se describe el procedimiento que lleva adelante el módulo para resolver una solicitud de tipo *ListRecords* a partir de la enunciación de 4 etapas sucesivas:

1. Inicio: A partir de la URL solicitada se determina el contexto deseado, el verbo OAI-PMH y un formato de metadatos, entre otros parámetros. Luego, en función de la configuración definida en *xoai.xml*, se determinan los *filter*, *transformers* y *metadataFormats* definidos para el contexto solicitado y que se utilizarán en las siguientes etapas.
2. Selección: Se invoca a cada uno de los filtros determinados en el paso anterior para definir las condiciones lógicas de selección de datos, que luego se combinarán para generar una consulta a la fuente de datos definida — *database o solr*—. Una vez generada dicha consulta, se ejecuta y se obtiene un conjunto de ítems del repositorio.
3. Transformación: Se procesan los ítems resultantes del paso anterior a partir del transformer definido en el paso 1. Los transformers pueden alterar los metadatos de cualquier forma, aunque deben mantener el formato XOAI[14] (ver Figura 5) de forma que, al continuar con el procesamiento en el paso 4, se conozca la estructura de ítems recibida.
4. Mapeo/crosswalk: esta fase consiste en tomar los ítems resultantes del paso anterior y aplicarles una nueva transformación XSLT para adecuar su estructura, actualmente en formato XOAI, al formato de metadatos solicitado.

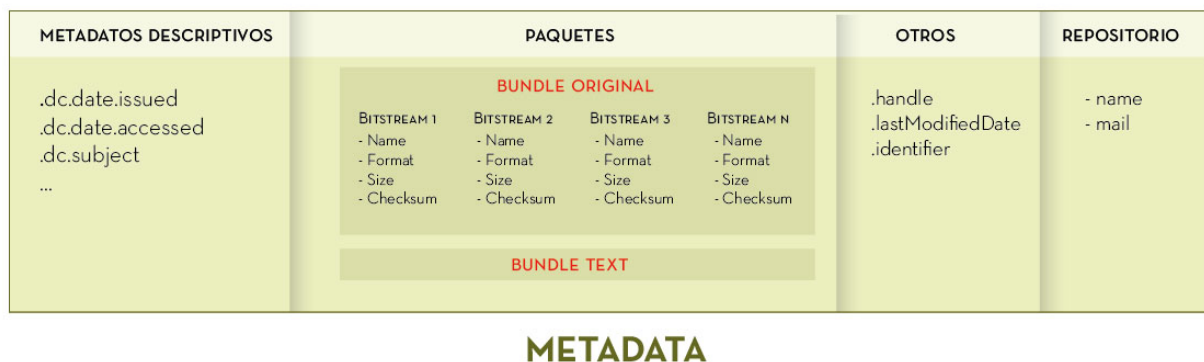


FIGURA 5 - Representación de metadatos XOAI.

Finalizada la última etapa, los metadatos son devueltos en la respuesta OAI-PMH. En caso de recibir una solicitud de verbo *GetRecord*, el procedimiento es prácticamente igual al planteado, con la siguiente excepción en el paso 2: en lugar de construir la consulta de recuperación de datos en función de los filtros, el módulo recupera el ítem y le pide luego a los filtros que lo validen. Si todos los filtros aprueban el ítem se procede con la etapa 3, caso contrario el procesamiento de la solicitud es abortado.

PERSONALIZACIÓN DEL MÓDULO OAI 2.0

Cabe destacar que el módulo OAI 2.0 de DSpace ya provee una definición para las directrices DRIVER y OpenAire, consistente en configuraciones genéricas que deben ser adaptadas a la estructura y necesidades de cada repositorio. De acuerdo a las características de los metadatos internos del repositorio será la complejidad y cantidad de filtros necesaria y la complejidad del transformer.

La configuración del *data provider* requiere tomar una serie de decisiones que incluye:

- I. Determinar a qué directrices el repositorio debe, o conviene, adecuarse. En general se querrá mantener configuraciones separadas por lo que para cada conjunto de directrices tendremos un contexto.
- II. Determinar qué ítems cumplen con las condiciones necesarias planteadas por las directrices y cuáles no. Por ejemplo, si se desean documentos de tipologías específicas o con acceso irrestricto, entre otros. Estas condiciones deberán ser traducidas en filtros que segreguen las obras deseadas por el contexto.
- III. Determinar qué requisitos estructurales plantean las directrices y analizar si es posible adecuar los metadatos internos a los mismos. Es normal que en las directrices se requieran metadatos que no existen en los ítems aunque, en general, es posible derivarlos a partir de otros metadatos pre-existentes. Estos cambios y derivaciones se plasman dentro de los transformers.
- IV. Determinar el esquema que debe mantener cada metadato de acuerdo a las restricciones de las directrices. Para algunos metadatos, las directrices requieren o aconsejan un esquema o dominio de datos válido. Algunos ejemplos son: iso639-2 para dc.language, tipología DRIVER para el campo dc.type, etc. Sin embargo y dado que internamente el repositorio guarda los metadatos bajo esquemas propios o diferentes al necesario, no queda otra opción que realizar un mapeo de esquemas donde se corresponda cada valor del repositorio con el esperado por las directrices. Al igual que el caso anterior, estos mapeos de esquemas se realizan desde el transformer.
- V. Determinar los formatos de metadatos necesarios en la disseminación. A partir de ésta decisión, se definirán los formatos del contexto.

Creación de contextos

En primer lugar se comienza con la declaración de cada contexto en el archivo *xoai.xml*, creando una etiqueta *Context*, con el atributo *baseurl* indicando el nombre público del contexto. Posteriormente en este elemento *Context*, haremos referencia a los otros componentes: *Filter*, *Format*, *Transformer* y *<Set*.

Es posible generar un *set* dentro de cada contexto que simplemente incluya todo el contenido del contexto de acuerdo a lo aceptado por sus filtros. Esta práctica es usuario y requerida en las 3 directrices ya mencionadas. En la Figura 6 se incluye un extracto de la definición de un contexto *snrd*, que retorna los recursos del repositorio que cumplen con las restricciones planteadas en las directrices SNRD.

```
<Context baseurl="snrd">  
  <Set refid="snrdSet"/>  
  <Filter refid="snrddocumentsubtypeFilter"/>  
  ...  
  <Filter refid="fulltextOnlyFilter"/>  
  <Filter refid="existsBundleOriginalFilter"/>  
  <Transformer refid="snrdTransformer"/>  
  <Formats refid="oaidc"/>  
</Context>
```

FIGURA 6 - Declaración de un contexto en OAI 2.0.

Creación de filtros (filters)

Para la definición de los filtros fue necesario analizar las directrices y detectar en cada caso qué tipologías de recursos eran aceptadas y cuáles eran las condiciones mínimas y obligatorias que debía cumplir. Luego, en cada caso se determinó si era necesario utilizar un filtro o si se podría resolver de otra forma, como por ejemplo desde un transformer.

La cadena de filtros a utilizar para cada contexto se especifica en la configuración del archivo *xoai.xml*, dentro de la declaración del contexto. Se pueden utilizar tanto filtros preexistentes —del paquete *org.dspace.xoai.filter*— como filtros propios, implementados en lenguaje Java © de acuerdo a como se explicará más adelante.

En general, los filtros aceptan algún valor en su configuración, dependiendo de los parámetros que éstos esperan para su funcionamiento. En la TABLA 3 se detallan algunos filtros predefinidos que seguramente se utilizarán en cualquier implementación:

Filtro	Descripción	Parámetros esperados
DSpaceAtLeastOneMetadataFilter	<p>Selecciona a partir del parámetro <i>field</i> los ítems que tienen dicho metadato válido de acuerdo los parámetros indicados en <i>operator</i> y <i>values</i>. Los operadores posibles son: CONTAINS, EQUAL, GREATER, LOWER, STARTS_WITH, GREATER_OR_EQUAL, LOWER_OR_EQUAL, ENDS_WITH. Por ejemplo si instanciamos este filtro con los parámetros <i>field=type</i>, <i>operator=EQUAL</i> y <i>values="Article, Book"</i> tendremos un filtro que acepta solamente artículos y libros.</p>	<p>- field - operator - values / value</p>
DSpaceMetadataExistsFilter	<p>Selecciona los ítems que contienen al menos una instancia de alguno de los metadatos indicados por el parámetro <i>fields</i>.</p>	<p>- field/fields</p>

TABLA 3 - Algunos filtros por defecto en OAI 2.0

En particular, los filtros mencionados en la tabla 3 proveen solamente acceso a los metadatos descriptivos del ítem. En cambio, si se desea filtrar a partir de otro criterio se deberá implementar un filtro ad-hoc, por ejemplo para hacer validaciones sobre bundles, bitstreams o visibilidad, entre otros.

A partir de DSpace 5.0 y OAI 2.1 también será posible generar filtros compuestos definidos como una combinación lógica de otros filtros: *AndFilter*, *OrFilter* y *NotFilter*. Estos filtros lógicos permitirán:

- reusar filtros ya definidos a partir de su combinación o negación
- crear nuevos filtros que seleccionen complementos. Es decir, en lugar de indicar por extensión todo lo que se quiere, se indica qué es lo que no se quiere.

Implementación de filtros nuevos

Cada filtro personalizado en OAI 2.0 debe incluir la lógica de validación estipulada por la clase *org.dspace.xoai.filter.DSpaceFilter*, teniendo en cuenta que cada método de filtrado se aplica sobre un modelo particular y en consecuencia de solicitudes OAI-PMH diferentes (ver Tabla 4). Se presenta entonces una situación muy propensa a errores que debe ser analizada con precaución al momento de configuración, especialmente cuando los filtros validen a partir de los bitstreams o de otros datos que NO sean sus metadatos descriptivos.

Verbos OAI-PMH	Método	Descripción	Modelo
ListRecords, ListIdentifiers	getWhere()	Devuelve una condición SQL a utilizar en la consulta sobre la base de datos usada por DSpace.	Base de datos
ListRecords, ListIdentifiers	getQuery()	Devuelve una condición SOLR a utilizar en la consulta sobre el motor de indexación.	Documento Solr
GetRecord	isShown()	Verifica si dado un ítem, el mismo es apto para ser incluido en el contexto.	DSpaceItem

TABLA 4 - Métodos a implementar sobre un filtro propio.

Se resume a continuación las características de cada uno de estos modelos:

1. Base de datos: es el modelo de entidades de DSpace, representado por sus tablas en la base de datos. Provee acceso irrestricto a todos los datos del repositorio.
2. Documento Solr: contiene los datos almacenados en el *core* de Solr correspondiente al módulo OAI2.0. Cada documento mantiene datos sobre un ítem, que provienen de la tabla *item*, y sus metadatos descriptivos, que provienen de la tabla *metadatatype*. Ver Figura 7. También contiene dos campos particulares:
 - *item.compile*: campo de cache que contiene un texto con el modelo XOAI completo del ítem serializado en XML, y
 - *metadata.dc.format.mimetype*: campo derivado que contiene el MimeType del primer bitstream del bundle ORIGINAL.
3. DSpaceItem: es el modelo XOAI mencionado previamente en la Figura 5, que contiene metadatos descriptivos, datos sobre bundles y otros datos sobre el ítem y el repositorio.

METADATA	ITEM	EXTRA
.dc.date.issued .dc.date.accessed .dc.subject ...	community collection lastModifiedDate Public Submitter	- mimeType (metadata.dc.format.mimetype) - item.compile

FIGURA 7 - Estructura de un documento Solr accesible desde el método *getQuery*



Obviar estas diferencias durante el armado de los filtros puede resultar en que los mismos no funcionen de acuerdo a lo esperado y/o que se comportan distinto en un `GetRecords` de un `ListRecords`. En particular

A modo de ejemplo se incluye en la Figura 8 la implementación de la clase `DspaceExistsBundleOriginalFilter` correspondiente al filtro 2.2 que se explicará posteriormente.

Creación de transformadores (transformers)

Luego de implementar los filtros se define si es necesario definir transformaciones sobre los metadatos. En caso de ser necesario alterar los ítems antes de pasar a la siguiente etapa de mapeo, se debe definir un archivo de transformación XSLT. Ésta hoja de estilos actuará sobre la estructura de metadatos XOAI (ejemplificada en la Figura 5), pudiendo inhabilitar la exportación de ciertos metadatos, transformar el valor de los existentes o generar algunos nuevos.

Para definir un transformer es necesario editar el archivo `xoai.xml` :

- declarar el nuevo transformer insertando un elemento *Transformer*, el cual consta de un atributo *id* arbitrario y un elemento subordinado *XSLT* que contiene una ruta relativa al archivo de transformación XSLT a utilizar dentro del directorio `transformers`.
- referenciar desde el contexto deseado al transformer recién declarado a partir de su atributo *id*.

```

public class DspaceExistsBundleOriginalFilter extends DSpaceFilter{

    public SolrFilterResult getQuery(){
        return new SolrFilterResult("metadata.dc.format.mimetype:[* TO *]");
    }

    public DatabaseFilterResult getWhere(Context context){
        try{
            return new DatabaseFilterResult("EXISTS (" +
                "SELECT distinct r.bundle_id "+
                "FROM " +
                "(SELECT a.bundle_id "+
                "FROM " +
                "(SELECT * "+
                "FROM item it INNER JOIN item2bundle i2b ON it.item_id = i2b.item_id" +
                "WHERE in_archive AND item_id = i.item_id) AS a " +
                "INNER JOIN "+
                "bundle b ON a.bundle_id = b.bundle_id" +
                "WHERE name = ?) as r " +
                "INNER JOIN "+
                "bundle2bitstream as b2b on r.bundle_id = b2b.bundle_id" +
                "ORDER BY r.bundle_id"+
                ")", "ORIGINAL");
        }catch (Exception e){
            log.error(e.getMessage(), e);
        }
        return new DatabaseFilterResult();
    }

    public boolean isShown(DSpaceItem item) {
        Iterator<Element> metadata = item.getMetadata().getMetadata().getElement().iterator();
        while(metadata.hasNext()){
            Element metadataElement = metadata.next();
            if (metadataElement.getName().equals("bundles")){
                Iterator<Element> bundles = metadataElement.getElement().iterator();
                while(bundles.hasNext()){
                    Element bundle = bundles.next();
                    if (bundle.getName().equals("bundle")){
                        Iterator<Field> fieldIter = bundle.getField().iterator();
                        while(fieldIter.hasNext()){
                            Field field = fieldIter.next();
                            if(field.getName().equals("name") && field.getValue().equals("ORIGINAL")){
                                return true;
                            }
                        }
                    }
                }
            }
        }
        return false;
    }
}

```

FIGURA 8 - Codificación de métodos para filtro particular.



Creación de mapeadores (metadataFormats)

Ésta es la etapa final realizada por el módulo antes de retornar la respuesta al verbo OAI-PMH *GetRecord* o *ListRecord* recibido. Los archivos utilizados aquí son similares a los archivos de transformación anteriores pero en lugar de modificar los datos y metadatos, se encargan principalmente de transformar los nombres de los metadatos en base a mapeos hacia un formato específico. Incluso reciben los metadatos bajo la misma estructura XOAI indicada en la Figura 5.

Por cada formato de metadatos declarado en el archivo *xoai.xml*, existe un archivo XSLT ubicado en el directorio *crosswalks/oai/metadataFormats/* que se encarga de realizar las correspondencias desde el formato interno de metadatos a un formato específico, como por ejemplo *oai_dc*[11], *mets* u otros. Cada mapeador de formato puede ser usado en cualquier contexto, agregando el elemento *Format* dentro de la declaración del elemento *Context* de *xoai.xml*.

En esta etapa se debe prestar particular atención a que los metadatos recibidos sean manejados explícitamente. Cualquier omisión de mapeo provocará que no se exponga algún metadato, afectando de forma negativa la calidad del recurso para los cosechadores.

Típicamente, cada metadato se maneja de las siguientes maneras:

- a) se exporta bajo el mismo nombre que trae. Caso típico: se usa internamente *dc.type* y se debe exponer *dc.type*
- b) se cambia el nombre del metadato interno por otro del formato destino. Ej. se pasa de *sedici.author.corporate* a *dc.creator*.
- c) Se combina 2 o más metadatos en uno solo. Ejemplo: se toma los campos internos *dc.title* y *sedici.subtitle* para generar el campo destino *dc.title* como la unión de los 2 anteriores separados por ': '.
- d) Se omite su exportación en caso que el metadato no sea relevante o no se desea que quede público. Ejemplo: *sedici.description.fulltext* es un campo que no debe ser exportado.

En cualquier caso pueden aplicarse transformaciones sobre los datos para respetar un esquema particular, como ISO 8601 para fechas y ISO 639-2 para idiomas, o para normalizar los textos, pasar un metadato a minúsculas, etc.

También resulta importante considerar el orden en que se exportan los metadatos. Por ejemplo, se suele asumir que la primer instancia del campo *dc.date* es la fecha que representa el recurso por lo que debe exportarse primero la fecha de publicación o presentación. Algo similar sucede con el título principal y títulos alternativos, autores y otros campos.



EXPERIENCIA EN EL REPOSITORIO SEDICI – UNLP

El repositorio SEDICI contaba originalmente con un único contexto global *por default*, denominado *request*, que respondía a todas las solicitudes OAI-PMH recibidas. Con el objetivo de dar cumplimiento a las directrices DRIVER, OpenAIRE y en particular SNRD, se configuraron y habilitaron los respectivos contextos.

Dado que los metadatos del repositorio SEDICI son bastante particulares y diferentes del Dublin Core tradicional, resultó necesario configurar varios filtros e implementar algunos otros. Se tiene en consecuencia que la solución encontrada difícilmente sea reusable tal y como está, dado que refleja una realidad muy particular.

Se enumeran a continuación sólo algunas condiciones que debieron ser consideradas sobre los ítems a retornar en cada contexto y que buscan ejemplificar los casos más generales:

- a) debe poseer un metadato *dc.type*
- b) la primer instancia del metadato *dc.type* debe tener una de tipología soportada por las directrices,
- c) la segunda instancia del metadato *dc.type* debe contener la versión,
- d) debe contener al menos un archivo a texto completo y accesible alojado en el repositorio,
- e) la primer instancia del metadato *dc.rights* debe indicar el nivel de acceso a la obra.

Para (a) no resultó necesario hacer nada dado que siempre se cumple que todo recurso del repositorio posee el metadato *dc.type*.

Para (b) debe tenerse en cuenta 2 cuestiones: b.1) que los tipos de documentos de SEDICI son propios y no coinciden con los de las directrices por lo que resulta necesario hacer una adaptación o mapeo desde los tipos internos a los de la directriz en cuestión, y b.2) que no todos los tipos internos pueden ser adaptados a los de las directrices y que por tanto no deben incluirse en el conjunto resultado. El punto b.1) debe resolverse dentro del transformer mientras que el b.2) requiere la aplicación de un filtro.

El caso (c) presenta una situación diferente ya que requiere la existencia de un metadato que no está presente en el formato interno del repositorio. Sin embargo, resultó posible derivar este campo de *versión* en función del tipo y subtipo de cada documento. Dado que se debía agregar un metadato al ítem, se determinó que era necesario agregar esta lógica al transformer correspondiente.



La accesibilidad de un recurso a texto completo para cada ítem (d) es una condición compleja de evaluar dado que en el repositorio se tienen recursos con archivos cargados en DSpace y recursos con referencias a archivos externos, para referencias a recursos alojados en repositorios externos. Además, se tiene también que en ciertos casos el recurso no está disponible a texto completo sino que solo se tiene un extracto del mismo. En consecuencia, resulta necesario definir un filtro que analice cada ítem y acepte sólo los recursos que tienen al menos un bitstream en el bundle ORIGINAL y que tienen un metadato que indique que está disponible a texto completo (sedici.description.fulltext en SEDICI)

Finalmente el último caso (e) implica 2 puntos: e.1) que se incluya en un metadato *dc.rights* el nivel de acceso y e.2) que dicha instancia se exponga la primera entre otras instancias de *dc.rights*. Ambas tareas implican la modificación del ítem por lo que deberán ser resueltas desde el transformer como sigue: para e.1) se debe derivar el campo en función de otros metadatos y bitstreams asociados al ítem, para e.2) deberá forzarse la salida de esta instancia antes que las otras instancias de *dc.rights*.

Filtros propios

De esta forma, podemos determinar la necesidad de 3 filtros (ver Figura 9):

1. Caso b.2: que acepte todos los documentos que tengan tipologías deseadas. Se puede implementar a partir de *DSpaceAtLeastOneMetadataFilter*
2. Caso d: que acepte sólo documentos con bitstreams en bundle ORIGINAL y metadato *sedici.description.fulltext* con valor 'true'. Se puede implementar usando 2 filtro en secuencia :
 1. un *DSpaceAtLeastOneMetadataFilter* que acepte documentos con metadato *sedici.description.fulltext* con valor 'true'
 2. un filtro *DSpaceExistsBundleOriginalFilter*, implementado ad-hoc, que valide la existencia de al menos un bitstream en el bundle ORIGINAL
 - 3.

El primer filtro se aplica directamente sobre el contexto *driver*, así como en *snrd* y *openaire*, ya que validan la mayoría de los aspectos de DRIVER. El segundo y tercer filtro son particularmente necesario en el contexto *snrd*.


```
<Filter id="driverDocumentTypeFilter">
  <Class>org.dspace.xoai.filter.DSpaceAtLeastOneMetadataFilter</Class>
  <Parameter key="field">
    <Value>sedici.subtype</Value>
  </Parameter>
  <Parameter key="operator">
    <Value>equal</Value>
  </Parameter>
  <Parameter key="value">
    <Value>Articulo</Value>
    <Value>Documento de trabajo</Value>
    <Value>Revision</Value>
    ...
  </Parameter>
</Filter>

<Filter id="fulltextOnlyFilter">
  <Class>org.dspace.xoai.filter.DSpaceAtLeastOneMetadataFilter</Class>
  <Parameter key="field">
    <Value>sedici.description.fulltext</Value>
  </Parameter>
  <Parameter key="operator">
    <Value>equal</Value>
  </Parameter>
  <Parameter key="value">
    <Value>>true</Value>
  </Parameter>
</Filter>

<Filter id="existsBundleOriginalFilter">
  <Class>ar.edu.unlp.sedici.dspace.xoai.filters.DspaceExistsBundleOriginalFilter</Class>
</Filter>
```

FIGURA 9 - Declaración de filtros en xoai.xml

Transformadores propios

De acuerdo con las restricciones mencionadas debe incluirse en implementación las siguientes 4 transformaciones:

1. Caso b.1) mapear tipología interna de SEDICI a la tipología de DRIVER,
2. Caso c) derivar la versión de cada obra en función del tipo y subtipo de cada documento
3. Forzar que la salida de T1 y T2 sean la primer y segunda instancia del metadato dc.type
4. Caso e.1) derivar el nivel de acceso de cada obra en función de otros metadatos y bitstreams asociados al ítem,
5. Caso e.2) la primer instancia del campo dc.rights debe contener el nivel de acceso según vocabulario OpenAIRE

En la Figura 10 se muestran fragmentos XSL que se corresponden con los transformers T1 y T3. No se incluye T2 y T4 dado que son similares a T1, ni T5 porque es similar a T3. Todas las transformaciones agregan nuevos elementos al árbol de metadatos XOAI, utilizando fragmentos de código XSL de transformación.

<pre> <!-- T1: Se calcula el tipo DRIVER --> <xsl:variable name="driverType"> <xsl:choose> <xsl:when test="\$sediciType='Documento de trabajo'"> info:eu-repo/semantics/workingPaper </xsl:when> <xsl:when test="\$sediciType='Preprint'"> info:eu-repo/semantics/preprint </xsl:when> <xsl:when test="\$sediciType='Articulo'"> info:eu-repo/semantics/article </xsl:when> ... <xsl:otherwise> info:eu-repo/semantics/other </xsl:otherwise> </xsl:choose> </xsl:variable> </pre>	<pre> <!-- T3: se produce dc.type ordenado --> <doc:element name="type"> <doc:element name="es"> <doc:field name="value"> <xsl:value-of select="normalize-space(\$driverType)"/> </doc:field> </doc:element> <doc:element name="es"> <doc:field name="value"> <xsl:value-of select="normalize-space(\$driverVersion)"/> </doc:field> </doc:element> <doc:element name="es"> <doc:field name="value"> <xsl:value-of select="normalize-space(\$sediciType)"/> </doc:field> </doc:element> </doc:element> </pre>
--	---

FIGURA 10 - Bloques XSLT de transformadores T1 y T3.

Como se indicó en la Tabla 1, cada directriz tiene su recomendación para el uso del metadato *dc.type*. Sin embargo todas las transformaciones de tipología tienen la misma estructura de mapeo o crosswalk entre vocabularios, es decir, que para cada tipo interno de los define una única correspondencia en el vocabulario destino. En el caso de SEDICI se debió definir para el primer transformer (T1) un mapeo que para cada tipo de SEDICI retorne uno DRIVER.

La versión a generar en el segundo transformer (T2) no forma parte del conjunto de metadatos almacenado sobre cada ítem, por lo que deberá generarse teniendo en cuenta la tipología y otras cualidades del material. Básicamente, es muy similar a lo realizado en el caso anterior, es decir, en función del valor de la tipología se determina la versión.

El tercer transformer (T3) tiene la particularidad de forzar la salida de T1 y T2 como primera y segunda instancia del metadato *dc.type*, seguida de todas las demás instancias existentes. Esta acción se realiza de forma simple en XSLT como se puede ver en la Figura 9, donde se imprime la primer instancia y luego se fuerza la impresión del resto.

El cuarto caso, T4, deriva el nivel de acceso a un ítem (abierto, cerrado, embargado, restringido) según la existencia de un embargo temporal, de archivos asociados al ítem, del nivel de acceso a los mismos y de la existencia de un metadato específico con una URL de acceso externo.

Finalmente, el último transformer (T5) es similar a T3, y hace que la primer instancia del metadato *dc.rights* contenga el nivel de acceso calculado en T4.

Formatos propios

Para la definición de los mapeos del repositorio SEDICI se utilizó como base el mapeo que trae consigo DSpace, que maneja los casos típicos de mapeo hacia los formatos más usados como oai_dc, mets y mods, entre otros. Particularmente para los contextos *driver*, *openaire* y *snrd* sólo se ajustó el archivo de mapeo *oai_dc.xsl* dado que dichas directrices requieren solo ese formato.

La personalización de *oai_dc.xsl* se hizo en función de una tabla de mapeos que define para cada metadato interno la correspondencia en el formato destino, es decir, Dublin Core simple. Como se mencionó previamente, las correspondencias pueden ser simples (1 a 1), compuestas (n a 1), o nulas (1 a 0), y pueden implicar o no transformaciones sobre los datos (ver Tabla 5).

#	Metadato origen	Metadato destino	Notas sobre el valor
1	dc.type	dc.type	no se altera
2	sedici.creator.corporate	dc.creator	no se altera
3	sedici.creator.person	dc.creator	no se altera
4	dc.title sedici.subtitle	dc.title	se combinan los 2 campos con ': '. Ver Figura 11
5	dc.title.alternative	dc.title	
6	dc.language	dc.language	Se adecua el esquema de idiomas a ISO 639-1

TABLA 5 - Mapeo de metadatos internos a Dublin Core Simple

```
<!-- dc.title=dc.title: sedici.subtitle -->
<xsl:if select="doc:metadata/doc:element[@name='dc']/doc:element[@name='title']/...">
  <dc:title>
    <xsl:value-of select="doc:metadata/doc:element[@name='dc']/doc:element[@name='title']/..." />
    <xsl:if test="position() = 1">
      <xsl:if test="doc:element[@name='sedici']/doc:element[@name='subtitle']/...">
        <xsl:text> : </xsl:text>
        <xsl:value-of select="doc:element[@name='sedici']/doc:element[@name='subtitle']/..." />
      </xsl:if>
    </xsl:if>
  </dc:title>
</xsl:if>

<!-- dc.title=dc.title.alternative -->
<xsl:for-each
select="doc:metadata/doc:element[@name='dc']/doc:element[@name='title']/doc:element[@name='alternative']...">
  <dc:title><xsl:value-of select="." /></dc:title>
</xsl:for-each>
```

FIGURA 11 - Mapeo para el campo *dc.title*

Un caso particular surge al generar el campo dc.format con el tipo MIME de la obra, dado que este no es un dato que se guarde junto a los metadatos descriptivos sino que el mismo está asociado a los bitstreams del item del bundle ORIGINAL. Por ello, la generación de este campo no implica un mapeo como los anteriores sino que requerirá acceder a otra sección del item recibido en formato XOAI donde se contiene datos sobre los bundles y bitstreams (ver Figura 12 y Figura 13)

```
<!-- bitstream.format -->
<xsl:for-each select="doc:metadata/doc:element[@name='bundles']/
  doc:element[@name='bundle' and doc:field[@name='name' and text()='ORIGINAL']]/
  doc:element[@name='bitstreams']/doc:element[@name='bitstream' and position()=1]/doc:field[@name='format']">
  <dc:format><xsl:value-of select="." /></dc:format>
</xsl:for-each>
```

FIGURA 12 - Generación del campo dc.format a partir del formato del primer bitstream del Bundle Original.

```
<element name="bundles">
  <element name="bundle">
    <field name="name">ORIGINAL</field>
    <element name="bitstreams">
      <element name="bitstream">
        <field name="name">portada-0001.pdf</field>
        <field name="format">application/pdf</field>
        <field name="size">28198</field>
        <field name="url">http://localhost:8080//bitstream/10915%2F1063/1/bitstream</field>
        <field name="checksum">c4e3648443bda78bc452783ef30f2bdf</field>
        <field name="checksumAlgorithm">MD5</field>
        <field name="sid">1</field>
      </element>
    </element>
  </element>
  ...
</element>
...
</element>
```

FIGURA 13 - Sección del modelo XOAI con el primer bitstream del Bundle Original.

CONCLUSIONES

Gracias a la flexibilidad provista por el módulo OAI 2.0 es posible generar diferentes vistas de un repositorio, tarea que no era posible previamente y que limitaba las vías de interoperabilidad para repositorios basados en el software DSpace en versiones previas a la 3.

El potencial de uso de este tipo de contextos o vistas personalizados es muy amplio y permite construir y exponer subconjuntos arbitrarios de recursos con diversos fines, como ser:



- cumplir con las restricciones de calidad y formato establecidos por una o más directrices de interoperabilidad
- compartir recursos con otros repositorios de materiales seleccionados. Por ejemplo se podría construir un contexto que exponga todos los recursos de una temática particular para que lo coseche un agregador temático o incluso de una tipología en particular para que los coseche un agregador de Tesis, Revistas o Libros.
- definir un mecanismo interno de control y análisis del repositorio, a partir del reporte en un contexto ad-hoc de los recursos que no cumplen con el nivel de calidad deseado por falta de metadatos obligatorios, archivos, etc.
- generar múltiples servidores OAI para permitir mayor diseminación en repositorios multi-institucionales.

Esta flexibilidad aportada por el módulo trae aparejada una mayor complejidad en la configuración, implementación, definición de mapeos y posibles errores que debe afrontar el encargado de interoperabilidad de un repositorio.

Finalmente, el análisis de las directrices y su contraste con el perfil interno de metadatos del repositorio resultará de vital importancia para determinar los requerimientos de cada caso. Como se indicó previamente, cada repositorio tendrá configuraciones muy particulares y dependientes de su realidad.

REFERENCIAS

- [1] DSpace 4.x Documentation - OAI 2.0 Server- DuraSpace Wiki. Revisado Julio, 27, 2014 desde <https://wiki.duraspace.org/display/DSDOC4x/OAI+2.0+Server>
- [2] Lyncode XOAI in Github - XOAI Wiki - (Revisado en Julio, 27, 2014). Disponible en: <https://github.com/lyncode/xoai/wiki>
- [3] SEDICI - Repositorio de la Universidad Nacional de La Plata - (Revisado en Julio, 27, 2014). Disponible en: <http://sedici.unlp.edu.ar/>
- [4] OpeAIRE_Guidelines: For_Literature_repositories [Directrices OpenAIRE: Para Repositorios Literarios]. - (Revisado en Julio, 27, 2014) - Disponible en: <https://guidelines.openaire.eu/>
- [5] OpenAIRE Guidelines v.1.1 - (Revisado en Julio, 28, 2014) - Disponible en: <https://www.openaire.eu>
- [6] Directrices DRIVER 2.0: Directrices para proveedores de contenido - Exponiendo recursos textuales mediante OAI-PMH] (Revisado en Julio, 28, 2014) - Disponible en: <http://www.driver-support.eu/>



- [7] Directrices SNRD: Directrices para proveedores de contenido del Sistema Nacional de Repositorios Digitales Ministerio de Ciencia, Tecnología e Innovación Productiva. (Revisado en Julio, 28, 2014) - Disponible en: http://repositorios.mincyt.gob.ar/pdfs/Directrices_SNRD_2012.pdf
- [8] The Open Archives Initiative Protocol for Metadata Harvesting (2008) - (Revisado en Julio, 28, 2014) - Disponible en: www.openarchives.org/OAI/openarchivesprotocol.html
- [9] Lyncode - XOAI Add-On for DSpace - (Revisado en Julio, 28, 2014) - Disponible en: <http://www.lyncode.com/dspace/addons/xoai>
- [10] XSLT Transformation (Version 1.0) - W3C Recommendation - (Revisado en Julio, 28, 2014) - Disponible en: <http://www.w3.org/TR/xslt>
- [11] OAI_DC XML Schema - (Revisado en Agosto, 10, 2014) - Disponible en: http://www.openarchives.org/OAI/2.0/oai_dc.xsd
- [12] DSpace Architecture - (Revisado en Agosto, 8, 2014) - <https://wiki.duraspace.org/display/DSDOC4x/Architecture>
- [13] Usage of Open Access Repository Software: Worldwide - OpenDoar 2014 - Revisado en Agosto, 8, 2014) - Disponible en: <http://www.opendoar.org/onechart.php>
- [14] XOAI XML Schema - Internal Schema - (Revisado en Agosto, 10, 2014) - Disponible en: <https://github.com/lyncode/xoai/blob/3.x/schemas/xoai.xsd>
- [15] El caso de Interoperabilidad para Repositorios de Acceso Abierto - (Revisado en Agosto, 14, 2014) - Disponible en: <https://www.coar-repositories.org>
- [16] Herramienta para la recolección de metadatos bibliográficos mediante el protocolo OAI-PMH - Introducción - Yusniel Hidalgo Delgado - (Revisado en Agosto, 14, 2014) - Disponible en: http://www.researchgate.net/publication/235934068_Herramienta_para_la_recoleccin_de_metadatos_bibliograficos_mediante_el_protocolo_OAI-PMH

LA INCORPORACIÓN DEL REGISTRO DE OBRAS Y AUTORES DE UNA INSTITUCIÓN: ESTUDIO DE CASO SEDICI-UNLP-SCOPUS

Jose Texier¹, Marisa R. De Giusti², Ariel Jorge Lira³, Gonzalo Luján Villarreal⁴

¹Universidad Nacional Experimental del Táchira (UNET), Venezuela. Servicio de Difusión de la Creación Intelectual (SEDICI), Universidad Nacional de La Plata, Argentina. jtexier@unet.edu.ve; dantexier@sedici.unlp.edu.ar

²Servicio de Difusión de la Creación Intelectual (SEDICI), Universidad Nacional de La Plata, Argentina y Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA).

³Servicio de Difusión de la Creación Intelectual (SEDICI), Universidad Nacional de La Plata, Argentina

⁴Servicio de Difusión de la Creación Intelectual (SEDICI), Universidad Nacional de La Plata, Argentina

RESUMEN

El Servicio de Difusión de la Creación Intelectual (SEDICI), es el repositorio institucional de la Universidad Nacional de La Plata (UNLP), cuenta con recursos de diferentes tipologías como artículos de revistas científicas, publicaciones en congresos, tesis (doctorado, maestría, especialización y pregrado), normativas y ordenanzas, libros y libros electrónicos, audios, fotografías, entre otros. El objetivo principal de este trabajo fue recuperar y dar cuenta de todas las obras producidas por autores de la UNLP, para dar difusión y visibilidad dentro de su repositorio institucional e incrementar las producciones presentes en el mismo. En julio del 2013 se inició una recopilación de todos los artículos, objetos de conferencia y artículos de revisión que se encuentran en la base de datos bibliográfica Scopus comprendidos entre 1927 y 2013. Para lograr tal objetivo, se definió un mecanismo semiautomático de incorporación de recursos desde Scopus al DSpace implementado en SEDICI. Este mecanismo fue desarrollado en sucesivas fases, que pueden resumirse en: recuperación desde Scopus, filtrado, limpieza, transformación y normalización de los metadatos, importación de autores y de los recursos, revisión, filtrado, incorporación de los PDF (si la licencia lo permite) y finalmente publicación en el repositorio. La consulta sobre Scopus finalizó el 1 de marzo de 2014 en la UNLP y fue posible gracias al servicio de pago que Elsevier brinda al Ministerio de Ciencia, Tecnología e Innovación Productiva de la República Argentina. Esto implica que este tipo de consultas puede ser realizada por otras instituciones públicas de investigación de la Argentina que estén bajo el paraguas del MinCyT. En resumen, el trabajo tuvo dos logros tangibles: se incorporaron a SEDICI 19.940 documentos con filiación de la UNLP y 89.903 autores de cuales al eliminar los no repetidos se redujo a 24.161 autores. Esta información permitió sumar recursos a la producción local alojada en el repositorio, mejorar la calidad de los metadatos ya almacenados e incrementar la visibilidad del repositorio y de la universidad. Por otro lado, a partir de ISSN/ESSN de los listados de revistas de DOAJ, PubMed, SherpaRomeo, SciELO y Redalyc, se relacionaron con los artículos importados desde Scopus, lo que permitió detectar que al menos un 74% de los documentos detectados admite su incorporación a texto completo en el repositorio, ya sea porque la revista es de acceso abierto, es una revista que permite preprint o postprint, o porque el autor ha pagado el arancel correspondiente para que su artículo esté en abierto.

Palabras clave: interoperabilidad, repositorios institucionales, SEDICI, DSpace, Scopus.

INTRODUCCIÓN

El Servicio de Difusión de la Creación Intelectual, SEDICI [1], es el repositorio institucional de la Universidad Nacional de La Plata (UNLP) que se encuentra implementado en DSpace [2]. Para agosto del 2014, cuenta con aproximadamente

32.000 recursos de diferentes tipologías como artículos de revistas científicas, publicaciones en congresos, tesis (doctorado, maestría, especialización y pregrado), normativas y ordenanzas, libros y libros electrónicos, audios, fotografías, entre otros [3]. Los repositorios institucionales son estructuras web interoperables encargadas de recopilar, catalogar, gestionar, acceder, difundir y preservar toda la producción académica y científica de las instituciones descriptos por medio de un conjunto de datos específicos (metadatos) [4]–[7].

A mediados del año 2013, se comenzó un estudio sobre la incorporación de los recursos (artículos, objetos de conferencia y artículos de revisión) desde la base de datos bibliográfica Scopus [8] al DSpace implementado en SEDICI, a partir de todas las filiaciones de la UNLP registradas en Scopus y comprendidos en el lapso

1927-2013. El objetivo del estudio fue recuperar y dar cuenta de todas las obras producidas por autores de la UNLP en Scopus (tiene aproximadamente de 22.000 revistas [9]), para dar difusión y visibilidad dentro de su repositorio institucional e incrementar las producciones presentes en el mismo.

Para lograr tal objetivo, se definió un mecanismo semiautomático de incorporación de recursos desde Scopus a SEDICI-DSpace que puede servir de modelo a otras universidades e instituciones. Este mecanismo fue desarrollado en sucesivas fases, que pueden resumirse en: recuperación desde Scopus, filtrado, limpieza, transformación y normalización de los metadatos, importación de autores y de los recursos, revisión, filtrado, incorporación de los PDF (si la licencia lo permite) y finalmente publicación en el repositorio. La consulta sobre Scopus finalizó el 1 de marzo de 2014 en la UNLP y fue posible gracias al servicio de pago que Elsevier (editorial que gestiona Scopus) brinda al Ministerio de Ciencia, Tecnología e Innovación Productiva de la República Argentina. Esto implica que este tipo de consultas puede ser realizada por otras instituciones públicas de investigación de la Argentina que estén bajo el paraguas de dicho Ministerio.

METODOLOGÍA

Las fases nombradas anteriormente se pueden resumir en 5 etapas:

Etapas 1 – Recuperación de metadatos

Consistió en consultar en Scopus los artículos, objetos de conferencia y artículos de revisión presentes en todas las dependencias relacionadas con la UNLP, y luego exportar los resultados obtenidos a archivos en formato CSV. La consulta y exportación finalizó en marzo del presente año, garantizando la recolección de la producción científica hasta el año 2013. Las 14 filiaciones y los artículos recopilados para cada una se observan en la Tabla 1.

Filiación	Art.
Observatorio Astronómico de La Plata	45
Centro de Química Inorgánica (CEQUINOR)	54
Instituto de Investigaciones Bioquímicas de La Plata (INIBIOLP)	98
Instituto de Botánica Carlos Spegazzini	134
Instituto de Física de Líquidos y Sistemas Biológicos (IFLYSIB)	140
Instituto de Limnología Dr. Raul A. Ringuelet	139
Centro de Investigación y Desarrollo en Criotecnología de Alimentos (CIDCA)	168
Centro de Investigación y Desarrollo En Ciencias Aplicadas Dr. Jorge Juan Ronco (CINDECA)	196
Centro de Investigaciones Geológicas - UNLP	194
Centro de Investigaciones Ópticas - La Plata	438
Instituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas INIFTA	864
UNLP Facultad de Ciencias Medicas	1,023
Museo de La Plata	2,420
UNLP	14,027
TOTAL	19,940

Tabla 1. Filiaciones y artículos en Scopus

Scopus, permite exportar una serie de metadatos que en su mayoría corresponden a parte del modelo de metadatos presente en SEDICI. El modelo de metadatos que se puede exportar de Scopus es:

- *Citation information*
 - Author(s) / Formato Apellidos coma Iniciales
 - Document title
 - Year
 - Source Title
 - Volume, Issue, Pages
 - Citation count
 - Source and Document Type
- *Bibliographical information*
 - Affiliations
 - Serial identifiers (e.g. ISSN)
 - DOI
 - PubMed ID
 - Publisher
 - Editor(s)
 - Language of Original Document
 - Correspondence Address
 - Abbreviated Source Title
- *Abstract and Keywords*
 - Abstract
 - Author Keywords
 - Index Keywords
- *Fund Details*
 - Number
 - Acronym
 - Sponsor

- *References*
 - References
- *Other information*
 - Tradenames and Manufacturers
 - Accession numbers and Chemicals
 - Conference information

Etapa 2 – Normalización

Se tomaron los siguientes metadatos de los artículos exportados por Scopus y se incorporaron a Google Refine [10], lo que permitió realizar análisis, transformaciones y normalizaciones de manera flexible y rápida en dicha herramienta.

- Author(s)
- Document title
- Year
- Source Title
- Volume
- Issue
- Pages start
- Pages end
- Pages count
- Cited by Citation count
- Link Scopus
- Affiliations
- Author with Affiliations
- Abstract
- Author Keywords
- Index Keywords
- References
- Correspondence Address
- Editors
- Sponsors
- Publisher
- Conference name
- Conference date
- Conference location
- Conference code
- ISSN
- ISBN
- CODEN
- DOI
- PubMed ID
- Language of Original Document
- Abbreviated Source Title
- Document Type
- Source

El total de registros incorporados a Google Refine desde los archivos CSV obtenidos desde Scopus fueron 19.940 artículos comprendidos entre 1920 y 2013. Luego, Google Refine, permite exportar la información en archivos con formatos XML de acuerdo con plantillas que el usuario puede definir. La plantilla que se usó para exportar los 19.940 artículos fue:

```
<?xml version="1.0" encoding="utf-8" standalone="no"?>
<dublin_core schema="dc">
<doc>
<field name="cd">{{(cells["Conference date"].value)}}</field>
<field name="y1">{{(jsonize(cells["Year"].value))}}</field>
<field name="ab"><![CDATA[{{(cells["Abstract"].value)}}]]></field>
<field name="p1">{{(jsonize(cells["Page count"].value))}}</field>
<field name="p2">{{(jsonize(cells["Page start"].value))}}-{{(jsonize(cells["Page end"].value))}}</field>
<field name="la">{{(cells["Language of Original Document"].value)}}</field>
<field name="t1"><![CDATA[{{(cells["Title"].value)}}]]></field>
<field name="pu"><![CDATA[{{(cells["Publisher"].value)}}]]></field>
<field name="ty">{{(cells["Document Type"].value)}}</field>
<field name="cl">{{(cells["Conference location"].value)}}</field>
<field name="li"><![CDATA[{{(cells["Link"].value)}}]]></field>
<field name="ur"><![CDATA[{{(cells["DOI"].value)}}]]></field>
<field name="in">{{(cells["ISSN"].value)}}</field>
<field name="ib">{{(cells["ISBN"].value)}}</field>
<field name="cn"><![CDATA[{{(cells["Conference name"].value)}}]]></field>
<field name="st"><![CDATA[{{(cells["Source title"].value)}}]]></field>
<field name="vo">{{(cells["Volume"].value)}}</field>
<field name="is">{{(cells["Issue"].value)}}</field>
<field name="pm">{{(cells["PubMed ID"].value)}}</field>
<field name="k1"><![CDATA[{{(join(cells["Index Keywords"].value.split("; "), ","))}}]]></field>
<field name="k1"><![CDATA[{{(join(cells["Index Keywords"].value.split("; "), ","))}}]]></field>
<field name="ka"><![CDATA[{{(join(cells["Author Keywords"].value.split("; "), ","))}}]]></field>
<field name="ka"><![CDATA[{{(join(cells["Author Keywords"].value.split("; "), ","))}}]]></field>
<field name="a1">{{(join(cells["Authors"].value.split(", "), ","))}}</field>
<field name="a1">{{(join(cells["Authors"].value.split(", "), ","))}}</field>
<field name="ed">{{(join(cells["Editors"].value.split(", "), ","))}}</field>
<field name="ed">{{(join(cells["Editors"].value.split(", "), ","))}}</field>
<field name="su">{{(cells["SubType"].value)}}</field>
</doc>
</dublin_core>
```

Luego de exportar esta plantilla en Google Refine, se obtuvo un archivo en formato XML que contiene los 19.940 registros, el cual sirvió como entrada a un programa en Java para transformarlo en otros archivos XML con las estructuras de los esquemas de metadatos DC, SEDICI y MODS que soporta el modelo de metadatos de SEDICI-DSpace.

Se destaca, que todos estos registros se relacionaron con los listados de revistas que entregan DOAJ [11], PubMed [12], SherpaRomeo [13], SciELO [14] y Redalyc [15]. Esta relación fue posible a través de los ISSN y ESN que tiene cada uno de los documentos importados de Scopus y de los listados de dichas revistas. Las categorías obtenidas como nuevos metadatos para cada documento de Scopus, con énfasis en el Acceso Abierto (Open Access) [16], [17], son:

- **CC:** artículos que están con licencia Creative Commons, gracias a la consulta a PubMed
- **DOAJ:** artículos que pertenecen a revistas en Acceso Abierto
- **OA:** artículos en Acceso Abierto en forma parcial o total, gracias a DOAJ
- **Scielo:** artículos que la revista a la que pertenece esta en SciELO
- **Redalyc:** artículos que la revista a la que pertenece esta en Redalyc
- **Verde¹:** artículos que la revista a la que pertenece esta en SherpaRomeo con el estatus de que se puede archivar el pre-print y el post-print o versión de editor/PDF
- **Amarillo¹:** artículos que la revista a la que pertenece esta en SherpaRomeo con el estatus que se puede archivar el pre-print (ie la versión previa a la revisión por pares)
- **Blanco¹:** artículos que la revista a la que pertenece esta en SherpaRomeo con el estatus donde el archivo no está formalmente admitido
- **Azul¹:** artículos que la revista a la que pertenece esta en SherpaRomeo con el estatus donde se puede archivar el post-print (ie la versión final posterior a la revisión por pares) o versión de editor/PDF
- **Otros:** artículos sin agrupamiento

Con esta clasificación, se agruparon los 19.940 registros que se encontraban en Google Refine de la siguiente manera (Tabla 2):

Clasificación por grupos	Ítems
CC	532
Content DOAJ	319
OA	738
Scielo	787
Redalyc	84
Sherpa - Verde	9,977
Sherpa - Amarillo	1,575
Sherpa - Blanco	974
Sherpa - Azul	201
Otros	4,753
TOTAL	19,940

Tabla 2. Número de ítems por metadata

Etapa 3 – Importación de autores

La base de datos de autores de SEDICI se encuentra en una base de datos secundaria en MySQL, la cual es accedida desde el módulo de *authorities* de DSpace cada vez que se debe seleccionar un autor. Cuando se estaba haciendo la prueba de importación, se tuvo que modificar el SEDICI-Dspace para que realizara una búsqueda en la base de datos secundaria.

Para cada autor de Scopus, se mantiene sus apellidos, iniciales de su nombre y una afiliación sin normalizar porque aparece en Scopus exactamente como el autor lo colocó en el documento que aprobó la revista. La importación de estos autores se realizó antes que los documentos en DSpace, garantizando la existencia de tales autores, los cuales están clasificados de la siguiente manera:

- Existen 84.903 autores en la base de datos recopilada de Scopus, que pueden estar repetidos.
- Ejecución de una operación para detectar duplicados en los 84.903, la lista de autores se redujo a 23.161.
- Existen 48.867 autores de los 84.903 exportados, que tienen en sus afiliaciones “UNLP” o “La Plata”, pero se aplicó una operación similar a la anterior para detectar duplicados, ese número se redujo a 7.538.
- Los 7.538 autores, se relacionaron con los existentes en la base de datos secundaria de autores SEDICI, y se obtuvieron 3.123 autores que no coinciden, es decir, son nuevos.
- Existen 26.554 (de los 84.903) autores que no tienen en sus afiliaciones “UNLP” o “La Plata”.

Debido a la diversidad de información de los autores, principalmente enfocado en su filiación, para cada artículo incorporado y aprobado por los catalogadores de SEDICI, los autores fueron revisados manualmente e incorporados a la base de datos secundaria de SEDICI, para luego ser detectados desde SEDICI-DSpace.

Etapas 4 – Importación a DSpace de registros

DSpace provee varios mecanismos de importación de recursos que varían en complejidad y completitud de la información que representan. El usado en este trabajo es el Simple Archive Format, el cual es un formato básico que consiste en que para cada documento se genera una carpeta (items) con un archivo textual de índice, los bitstreams y con los metadatos expresados en archivos XML separados en archivos por cada formato de metadatos (DC, MODS & SEDICI) de acuerdo con el modelo de metadatos de SEDICI-DSpace. Un ejemplo de la operación a nivel de consola es:

```
/dspace import -a -e dantexier@hotmail.com -w -c 10915/25294 -s items/ -m mapfile.in
```

Se realizó la importación de prueba en el servidor de pruebas y de desarrollo sin problemas de 19.940 registros en 2241 seg. Después se actualizó el *discovery* para que DSpace reconozca los nuevos registros ingresados por la importación. Esta importación se realizó en una computadora portátil Sony VAIO con las siguientes características:

- Procesador: Intel Core i5-3337U CPU @ 1.80 GHz x 4
- Memoria RAM: 8 GB
- Disco duro de estado sólido (SSD) de 160 GB.
- Sistema operativo: Ubuntu 12.04 (precise) 64 bits y un kernel linux 3.5.0-23.generic

Otra opción para importar items en DSpace es Biblio-Transformation-Engine (BTE), pero las pruebas que se realizaron y por disponibilidad de recursos humano, se decidió continuar el proceso con la opción anterior¹.

Antes de realizar la importación, se tomo en cuenta el proceso de catalogación final realizado por los bibliotecarios de SEDICI, porque son ellos quienes realizan la revisión de cada unos de los metadatos importados desde Scopus, los normalizan, descargan los artículos desde los sitios web de las revistas (siempre y cuando la licencia lo permita) y asignan el licenciamiento de acuerdo con las políticas de cada una de las revistas de donde proviene el ítem. Por tanto, se agruparon los artículos en colecciones. Las colecciones, con base en las relaciones establecidas con las revistas comentadas anteriormente, que se crearon son:

- CC
- DOAJ
- OA
- Scielo
- SherpaRomeo
- Redalyc
- Otros

Etapas 5 – Deduplicación

Luego de la importación de autores y de los documentos, se hace una revisión y filtrado de cada uno, gracias a una tarea de curation para detectar duplicados, después si la licencia lo permite se incorporan los PDF y finalmente se autoriza la publicación en el repositorio por parte de los catalogadores. La tarea de curation se implemento usando la interfaz AbstractCurationTask del paquete org.dspace.curate. En esta interfaz se establece una conexión con SOLR a través de las clases CommonsHttpSolrServer, ModifiableSolrParams, QueryResponse y SolrDocumentList. Pero antes, se debe configurar tres SearchHandler para hacer búsquedas por los campos dc.title, dc.description.abstract y sedici.creator.person. De esta manera, se filtra la búsqueda para detectar el duplicado por la intercepción obtenidos por cada SearchHandler y configurando el parámetro "mm". A continuación parte del código del archivo de configuración:

¹ Info en la Wiki de DSpace: <https://wiki.duraspace.org/pages/viewpage.action?pageId=32481931>


```
<requestHandler name="titulos" class="solr.SearchHandler">
  <lst name="defaults">
    <str name="echoParams">explicit</str>
    <int name="rows">10</int>
    <str name="defType">edismax</str>
      <str name="q.alt">*:*</str>
      <str name="q.op">AND</str>
    <float name="tie">0.01</float>
    <str name="qf">
      dc.title^0.9
    </str>
    <str name="mm">1<&lt;75%</str>
  </lst>
</requestHandler>
```

Luego de obtener, vía SOLR, las tres consultas de acuerdo con el título, resumen y autores, se realiza una intercepción entre ellos cuyo resultado son ítems duplicados.

RESULTADOS

Los documentos incorporados en SEDICI, permitieron iniciar los procesos de eliminación de duplicados en todos los documentos del repositorio, mejoraron la calidad de los metadatos y se encontraron PDF que algunos registros existentes en SEDICI no los tenía. Además, esta información permitió sumar recursos a la producción local alojada e incrementar la visibilidad del repositorio y de la universidad.

En resumen, el trabajo tuvo dos logros tangibles: se incorporaron a SEDICI 19.940 documentos con filiación de la UNLP y 89.903 autores de cuales al eliminar los no repetidos se redujo a 24.161 autores de los cuales 3.123 son nuevos autores en SEDICI. Por otro lado, este estudio permitió detectar que al menos un 74% (Figura 1) de los documentos detectados de acuerdo con la licencia de la revista a la cual pertenece, admite su incorporación a texto completo en el repositorio, ya sea porque la revista es de Acceso Abierto, es una revista que permite preprint o postprint, o porque el autor ha pagado el arancel correspondiente para que su artículo esté en abierto. En la Tabla 3, se exponen los artículos importados desde Scopus que son candidatos a estar en Acceso Abierto y clasificados por año.

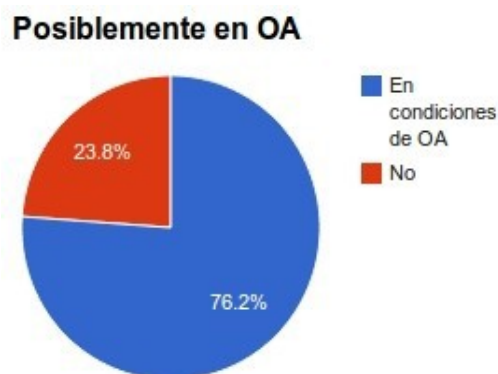


Figura 1. Artículos importados candidatos a estar en Acceso Abierto



No OA y OA por año		
Año	OA	No OA
1,920	21	1
1,927	0	1
1,929	2	0
1,930	0	1
1,932	3	0
1,934	1	0
1,939	1	0
1,944	1	1
1,945	1	0
1,946	1	0
1,947	0	1
1,949	1	0
1,952	1	0
1,953	1	0
1,954	0	1
1,955	1	1
1,956	2	2
1,958	0	1
1,959	1	0
1,960	2	0
1,961	2	1
1,962	12	3
1,963	5	8
1,964	11	3
1,965	13	8
1,966	12	6
1,967	23	3
1,968	21	7
1,969	23	6
1,970	40	7
1,971	57	12
1,972	43	12
1,973	55	30
1,974	52	26
1,975	38	29
1,976	39	26
1,977	36	33
1,978	34	24
1,979	44	22
1,980	57	23
1,981	59	39
1,982	75	43
1,983	76	47
1,984	84	62
1,985	109	47
1,986	105	46
1,987	136	69
1,988	145	69
1,989	142	64
1,990	164	68
1,991	160	70
1,992	157	57
1,993	173	75
1,994	205	108
1,995	217	81
1,996	413	119
1,997	433	145
1,998	460	145

1,999	465	153
2,000	512	140
2,001	552	179
2,002	546	168
2,003	594	149
2,004	592	173
2,005	641	200
2,006	662	225
2,007	745	214
2,008	809	264
2,009	934	234
2,010	994	246
2,011	964	240
2,012	1,013	262
2,013	1,194	253
Total	15,187	4,753

Tabla 3. Número de artículos incorporados

CONSIDERACIONES FINALES

El trabajo realizado, cumplió con el objetivo principal propuesto de recuperar y dar cuenta de todas las obras producidas por autores de la UNLP, además ayudo a incrementar la visibilidad del repositorio institucional SEDICI. La implementación de esta propuesta fue extensa y con una diversidad de opciones tecnológicas, la cual podrá mejorarse, por esta razón todo el código utilizado esta compartido en un proyecto GitHub [18].

De igual manera, se recomienda realizar la tarea de curation para detectar duplicados de forma periódica, ya que el deseo es tener la mejor calidad posible de metadatos del repositorio SEDICI.

Finalmente, se recomienda las siguientes líneas de trabajo en pro de consolidar el objeto principal de esta investigación, ofrecer e incrementar las diferentes obras de la UNLP, y mejorar la visibilidad del repositorio:

- Simplificar el proceso de incorporación de Scopus a cualquier repositorio a través de un framework que permita la representación de dichos recursos de forma independiente, para luego adaptarlos a SEDICI o a cualquier otro repositorio.
- Replicar este estudio en otros repositorios y/o otras plataformas de software.
- Mejorar la tarea de curation deduplicación y compararla con otras alternativas de detección de duplicados.



REFERENCIAS

- [1] M. De Giusti, N. Oviedo, A. Lira, A. Sobrado, J. Martinez, and A. Pinto, "SEDICI – Desafíos y experiencias en la vida de un repositorio digital," *RENATA*, vol. 1, no. 2, pp. 16–33, Aug. 2011.
- [2] DSpace, "DSpace," 2014. [Online]. Available: <http://www.dspace.org/>. [Accessed: 12-Mar-2014].
- [3] M. R. De Giusti, A. Sobrado, A. J. Lira, M. M. Vila, and G. L. Villarreal, "SEDICI (Servicio de Difusión de la Creación Intelectual)," *D-Lib Magazine*, vol. 14, Sep. 2008.
- [4] J. Texier, "Los repositorios institucionales y las bibliotecas digitales: una somera revisión bibliográfica y su relación en la educación superior," at the 11th Latin American and Caribbean Conference for Engineering and Technology - 2013, Cancun, Mexico, 2013.
- [5] J. Tramullas, "Bibliotecas digitales," presented at the VI Seminario de Centros de Documentación Ambiental y Espacios Naturales Protegidos, 2007.
- [6] C. L. Liew, "Digital library research 1997-2007: Organisational and people issues," *Journal of Documentation*, vol. 65, no. 2, pp. 245–266, Jun. 2009.
- [7] I. H. Witten, D. Bainbridge, and D. M. Nichols, *How to Build a Digital Library*. Morgan Kaufmann, 2003.
- [8] Elsevier, "Scopus – Welcome to Scopus," 2014. [Online]. Available: <http://www.scopus.com/>. [Accessed: 14-Aug-2014].
- [9] Elsevier, "Scopus | Elsevier," 2014. [Online]. Available: <http://www.elsevier.com/online-tools/scopus>. [Accessed: 15-Aug-2014].
- [10] OpenRefine, "OpenRefine," 2014. [Online]. Available: <http://openrefine.org/>. [Accessed: 28-May-2014].
- [11] DOAJ, "Directory of Open Access Journals," 2014. [Online]. Available: <http://doaj.org/>. [Accessed: 15-Aug-2014].
- [12] NCBI, "Home – PubMed - NCBI." [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed>. [Accessed: 15-Aug-2014].
- [13] JISC, "SHERPA/RoMEO," 2014. [Online]. Available: <http://www.sherpa.ac.uk/romeo/?la=es>. [Accessed: 15-Aug-2014].
- [14] SciELO, "SciELO - Scientific Electronic Library Online," 2014. [Online]. Available: <http://www.scielo.org/php/index.php?lang=es>. [Accessed: 15-Aug-2014].
- [15] Redalyc, "Redalyc. Latin America, the Caribbean, Spain and Portugal Scientific Journals Network.," 2014. [Online]. Available: <http://www.redalyc.org/home.oa>. [Accessed: 15-Aug-2014].



[16]B.-C. Björk and D. Solomon, "Open access versus subscription journals: a comparison of scientific impact," *BMC Medicine*, vol. 10, no. 1, p. 73, Jul. 2012.

[17]P. Suber, "Timeline of the Open Access Movement," 2009. [Online]. Available: <http://www.earlham.edu/~peters/fos/timeline.htm>. [Accessed: 31-Jan-2014].

[18]J. Texier, "ReviewRR- GitHub," 2014. [Online]. Available: <https://github.com/dantexier/ReviewRR/>. [Accessed: 06-Nov-2014].



INTEROPERABILIDAD ENTRE SISTEMAS DEL INSTITUTO ESPAÑOL DE OCEANOGRAFÍA

Concha Mosquera-de-Arancibia¹, Sergio Nieto-Caramés²

¹ Oceanógrafa coordinadora del Repositorio Institucional de Acceso Abierto del Instituto Español de Oceanografía. ² Ingeniero Informático responsable de Desarrollos Dspace en la empresa Arvo Consultores y Tecnología SL

RESUMEN

Aun cuando habitualmente en las primeras fases del ciclo de implantación de un RI (repositorio institucional), este se considera como un sistema aislado, inmediatamente, aparecen necesidades para que el RI interopere con el resto de sistemas de la organización, pasando a formar parte de lo que podríamos denominar el ecosistema de investigación institucional. Como elementos de esta necesaria interoperabilidad sistémica, destacaríamos, como pieza clave para la simplificación del sistema de trabajo ofrecido a los investigadores de recursos marinos y oceanografía, la integración de e-IEO (Repositorio Institucional Digital de Acceso Abierto del Instituto Español de Oceanografía) con SIPI (Sistema de Seguimiento Integrado de Proyectos de Investigación), que forma parte del Sistema de Gestión de Investigación del IEO. Se presentan los diversos elementos que el Instituto Español de Oceanografía está incorporando a su repositorio para lograr, de forma efectiva, la integración con el resto de sistemas de la organización, entre los que destacamos: 1. Integración de DSpace y SIPI, describiendo los elementos tecnológicos, procedimentales y organizativos necesarios para lograr la interoperabilidad semántica y técnica entre ambos sistemas y evitar a los investigadores el doble archivo. 2. Estandarización de los nombres de autor, mediante la implantación del control de autoridades (authority control) de DSpace, haciendo hincapié en el modelo elegido en e-IEO, en los requerimientos para aplicar esta mejora y en las ventajas obtenidas. 3. Adopción de identificadores ORCID (Open Researcher and Contributor ID) para la identificación y detección de ambigüedades y duplicidades en los nombres de los investigadores, usando un identificador único de autor de amplia difusión y uso. ORCID está conectado a otros sistemas actuales de identificación de autor como Author Resolver, Inspire, IraLIS, Scopus Author Identifier y otros. 4. Inclusión de vocabularios controlados específicos de ciencias marinas (recursos marinos, pesquerías, etc.) como mecanismo de mejora de las capacidades de búsqueda y descubrimiento de los ítems, y para posibilitar la interoperabilidad semántica e incrementar la visibilidad de los ítems en repositorios y recolectores.

Palabras clave: Instituto Español de Oceanografía, repositorio institucional digital, DSpace, interoperabilidad, control de autoridades, identificadores únicos de autor, vocabularios controlados.



ABSTRACT

Interoperability between systems of the Spanish Institute of Oceanography

Even though in the early stages of the cycle of implementation of an IR (institutional repository), this is considered as an isolated system, immediately, are needs that the IR interoperate with other systems in the Organization, becoming part of what we might call the ecosystem of institutional research. As this required systemic interoperability elements, would emphasize, as key for the simplification of the system of work offered to researchers of Oceanography and Marine Resources, the integration of e-IEO (Spanish Institute of Oceanography digital institutional open access repository) with SIPI (database of monitoring integrated research projects), which is part of the system of management of research of the IEO. The various elements that the Spanish Institute of Oceanography is incorporating into your repository to achieve effective integration with other systems of the Organization, among which we highlight are: 1. Integration of DSpace and SIPI, describing the technological, procedural and organizational elements required to achieve semantic and technical interoperability between the two systems and the researchers avoid the double file. 2. Standardization of the names of the author, through the implementation of authority control from DSpace, emphasizing the model chosen in e-IEO, the requirements to implement this improvement and the gains. 3. Adoption of identifiers ORCID (Open Researcher and Contributor ID) for the identification and detection of ambiguities and duplications in the names of the researchers, using a unique identifier of author of wide dissemination and use. ORCID is connected to other current author identification systems as Author Resolver, Inspire, Iralis, Scopus Author Identifier and others. 4. Inclusion of specific controlled vocabularies of marine sciences (marine resources, fisheries, etc.) as a mechanism for improving capabilities of search and discovery of the items, and to enable semantic interoperability and increase the visibility of items in repositories and harvesters.

Keywords: Spanish Institute of Oceanography, digital institutional repository, DSpace, interoperability, authority control, author identifiers, controlled vocabularies.



INTRODUCCIÓN

En esta comunicación se presentan los sistemas de información incorporados al ecosistema de investigación institucional del IEO, considerado este como el conjunto de elementos de apoyo, soporte y difusión de la actividad investigadora del Instituto, y en el que destacaríamos como elementos principales los siguientes:

- Repositorio Institucional Digital de Acceso Abierto del Instituto Español de Oceanografía, e-IEO.
- Sistema de Seguimiento Integrado de Proyectos de Investigación, SIPI.
- Base de Datos de Autoridades de Autor, BDAA.

La interoperabilidad sistémica que se construye entre estos sistemas, producto de actuaciones en los ejes tecnológicos, procedimentales y organizativos, es vital para lograr una consistencia efectiva entre los datos manejados e intercambiados, con el objetivo final de valorizar la producción investigadora de la Institución y de sus investigadores.

INTEROPERABILIDAD CON EL SISTEMA DE SEGUIMIENTO INTEGRADO DE PROYECTOS DE INVESTIGACIÓN

Desde mediados del año 2014, se han separado funcional y operativamente los sistemas de gestión de resultados de investigación del IEO. Por una parte el sistema de Seguimiento Integrado de Proyectos de Investigación, SIPI, se usa para la gestión interna de los resultados de investigación del IEO, mientras que el repositorio e-IEO, es usado para la ingesta, revisión y exposición de los objetos de investigación.

Esta separación operativa evita la doble descripción de los objetos digitales en SIPI y e-IEO, inasumible por los investigadores, pero en cambio obliga a sincronizar los datos de DSpace en la plataforma SIPI. Este alineamiento, que básicamente consiste en lograr la interoperabilidad semántica entre ambos sistemas, se logró mediante tres mecanismos principales:

- Alinear los elementos de metadatos descriptivos
- Alinear los valores de metadatos
- Conectar y sincronizar ambos sistemas, e-IEO y SIPI

El primer punto se abordó mediante la inclusión de más de 30 elementos nuevos al esquema Dublin Core Cualificado usado por DSpace (que está derivado del Library Application Profile del DCMI-Libraries Working Group). Se escogió realizar una extensión sobre este esquema frente a la opción de añadir un nuevo esquema específico del proyecto, principalmente por la dificultad de gestionar adecuadamente un nuevo esquema en los procesos de ingesta y exportación masivos.



Este crecimiento de elementos de metadatos implicó a su vez un proceso de adaptación de los formularios de entrada de datos y de visualización de ítems en DSpace, para dar cabida a esos 30 elementos añadidos.

El segundo reto era conseguir adaptar entre ambos sistemas los valores admisibles de los elementos de metadatos utilizados. Entre los elementos tecnológicos que se han incorporado de forma amplia a DSpace tenemos los esquemas semánticos, y bajo este término incluimos vocabularios controlados, tesauros y listados de encabezamiento de materias. Señalar, para dar una idea de la adaptación requerida, que en la versión actual de e-IEO, 17 elementos de metadatos incorporan alguno de estos esquemas semánticos. Éstos son principalmente listas de valores controlados, 11 elementos, pero incluyen también valores controlados de autoridad, 5 elementos, que se detallan en apartado siguiente y un elemento con un vocabulario controlado temático asociado.

Por último, como modelo de conexión y sincronización del e-IEO con el Sistema SIPI, se ha optado por la integración a través ficheros de intercambio, de los contenidos nuevos o actualizados de DSpace. De forma periódica, DSpace pone a disposición de SIPI la información sobre los ítems nuevos o modificados que hayan aparecido en DSpace, en los requeridos formatos de intercambio, con el fin de que SIPI pueda incorporar la información descriptiva asociada a los objetos digitales.

El repositorio e-IEO genera un fichero de texto, en un formato de intercambio definido conjuntamente entre los equipos responsables de ambos sistemas, cada vez que se detecta un cambio en alguno de los objetos almacenados. SIPI, al recibir esos ficheros, los procesa e incorpora así estos cambios a sus procesos.

Para la generación automática de los ficheros de intercambio se ha añadido un proceso a la cola de eventos de DSpace y configurado para generar un evento de un tipo cambio de ítem cada vez que se produce una creación, modificación o eliminación de un metadato existente o de un objeto. Estos eventos se añaden a la cola de eventos que es procesada cada vez que un evento ocurre para evitar sobrecarga en el sistema. Cada evento procesado genera un fichero con un nombre estructurado del tipo "AAMMDDHHMMSSTIPOITEM.txt". Esta nomenclatura simplifica el procesado de los ficheros por SIPI, pues así, cambios sucesivos de un mismo ítem de DSpace se ejecutarán también en el orden cronológico correcto en SIPI y evitaremos la posible inconsistencia de los metadatos.



CONTROL DE AUTORIDADES

El e-IEO usa el modelo de autoridades de DSpace (plugin de DSpace, Authority Control) para la validación interna de sus autores, implementado este sobre el metadato dc.contributor.author, principalmente, aunque se extiende al resto de elementos de la tipología dc.contributor.x.

Este metadato está conectado con uno los registros de autores del IEO, en forma de Base de Datos de Autoridades de Autor, BDAA, que contiene los nombres de los autores de la institución y los códigos de autoridades, entre otros datos. En el caso del e-IEO los registros de autoridad, contienen la forma autorizada del nombre del autor, establecida por la normativa DRIVER (DRIVER 2.0., 2008) como forma preferida, así, por ejemplo, el autor José Francisco Domínguez Yanes, se describe en la BDAA, y por ende en el repositorio e-IEO, según su nombre normalizado DRIVER, Domínguez-Yanes, J.F. (José Francisco).

Los objetivos del modelo de autoridades implantado en el e-IEO son los siguientes:

- Dar consistencia e integridad a los metadatos, ayudando en la corrección de los correspondientes valores. En el repositorio e-IEO se ha conseguido conectando las interfaces de archivo y edición de DSpace con una base de datos de autores del IEO con el fin de chequear los valores introducidos contra los registros de autoridad y poderles asignar una clave de autoridad única.
- Conseguir mejorar la precisión en la recuperación de la información, puesto que el mejor método, simple y positivo, de determinar si dos valores son idénticos, es comparando las claves de autoridad, ya que comparar valores textuales proporciona falsos positivos (demasiados García, M.) o falsos negativos (¿García, M. vs. García, Manuel?)
- Facilitar el intercambio de información bibliográfica con el SIPI, puesto que en la transferencia de información del e-IEO a SIPI se envían los valores de clave de autoridad, reconocibles y comunes entre ambos sistemas, en vez de los valores textuales de los nombres de autor.

En el e-IEO se usa, además, el modelo de *valores de confianza (confidence values)* de DSpace para mejorar la operativa del sistema de validación de autores. El valor de confianza se asigna adicionalmente al valor de clave de autoridad y se expresa como un valor simbólico dentro del rango *aceptado, incierto, ambiguo, no encontrado y sin validar*. Los valores de confianza aplicados en el e-IEO son los siguientes:

- *Aceptado*: el autor ha sido validado por un usuario (con permisos de envío, edición o revisión)
- *Incierto*: el autor ha sido validado por los procesos automatizados de validación de autores. En los procesos del repositorio, este valor de confianza tiene el mismo uso que el valor de aceptado.
- *Ambiguo*: se han encontrado varias posibles coincidencias en la BDAA, por lo que el usuario ha de desambiguar el valor en un proceso de edición manual.
- *No encontrado*: el autor no pertenece a la BDAA.
- *Sin validar*: el autor está pendiente de validar.

El modelo de Authority Control se aplica en el e-IEO por medio de varios procesos o en diversos momentos del ciclo de vida de un ítem:

- Proceso de depósito de ítems. Mediante la validación manual al insertar un nuevo ítem en el repositorio. Consiste en validar los autores mediante una pantalla de validación en el formulario de envío. Si se valida un autor contra la BDAA, a dicho autor se le normaliza el nombre al valor almacenado en la BDAA, se le asigna una clave de autoridad y se le asigna un valor de confianza de *aceptado*.
- Proceso de edición/revisión de ítems. Para los ítems archivados que requieran la modificación de los metadatos controlados por autoridad, `dc.contributor.author`, se puede aplicar el mismo proceso que en el depósito de ítems. Mediante una ventana el editor/revisor puede visualizar todos los valores de autor que coincidan parcial o totalmente con el nombre almacenado. Al seleccionar uno, se efectúa la validación del autor, con las acciones anteriormente citadas: normalización de nombre, asignación de clave de autoridad y asignación de valor de confianza.
- Tareas de curación. Debido al alto número de ítems que requieren validación de autores, ya que existen procesos de carga masiva de objetos provenientes de sistemas externos, en el e-IEO se implantó un tarea de curación que efectúa la validación automática de autores contra la BDAA. Este proceso se explica en los párrafos siguientes.

Para facilitar la validación de autores, se han implementado procesos de curación, procesos que pueden ejecutarse para comunidades, colecciones o ítems aislados. Estos procesos validan el metadato `dc.contributor.author` contra la BDAA. De entre los métodos disponibles en el Authority Control de DSpace se ha escogido el `getBestMatch` (lograr la mejor coincidencia) que confronta el valor del metadato 'nombre y apellidos del autor' con la BDAA, con los siguientes resultados:



- Si hay una única coincidencia con la BDAA, se asigna la clave de autoridad correspondiente, se normaliza el nombre y se asigna un valor de confianza de *incierto* (puede ser que el proceso se ha realizado sin intervención humana, no se asigna al valor *aceptado*, aunque son valores de confianza funcionalmente equivalentes)
- Si hay más de una coincidencia, no se asigna clave de autoridad ni se normaliza el nombre, pero se le asigna un valor de confianza de *ambiguo*, indicando que un usuario deberá de desambiguar manualmente el nombre de autor.
- Si no hay coincidencias, se asigna un valor de confianza de *no encontrado*.

IDENTIFICADORES ÚNICOS DE AUTOR

ORCID es un registro abierto de identificadores de investigadores y autores, con enlace a sus publicaciones, que nace como servicio en el año 2012. Pretende convertirse en un estándar para la identificación única y persistente de autores.

Aunque inicialmente su crecimiento se sustentó en el registro individual por los propios autores (en febrero de 2013 tenía 60.000 registros de autor), en la actualidad sus cifras parecen derivar del registro de Instituciones, Editores, Proveedores de servicios de Información, Integradores de software, etc. (110 miembros en julio 2014) y la consiguiente afiliación automática de sus autores (en julio de 2014, 800.000 registros).

Aunque el IEO participa en e-Ciencia (proyecto enmarcado en el convenio de cooperación interbibliotecaria entre la Comunidad de Madrid y el Consorcio Madroño para crear una plataforma digital de acceso abierto a la producción científica de la Comunidad de Madrid) y Madroño, Consorcio de Universidades de la Comunidad de Madrid para la Cooperación Bibliotecaria, adherido a ORCID, aún no ha efectuado un proceso de afiliación masiva de sus aproximadamente 700 investigadores y autores. No obstante, un número indeterminado de estos ha realizado su registro en ORCID y IEO ha considerado necesario empezar a incorporar funcionalidades de integración con la plataforma orcid.org para la validación internacional de sus autores.

En esta primera fase, el identificador ORCID, para los autores que dispongan del mismo y que lo notifiquen a los responsables del repositorio, se ha incluido en la BDAA, lo que posibilita que dichos identificadores se puedan usar para ofrecer servicios de conexión entre sistemas. Esto se consigue mediante la creación de relaciones de enlace entre los valores de autoridad locales, derivados de la BDAA y los valores de autoridad externos de ORCID. El IEO considera que al exponer sus autores validados a los usuarios finales, estas relaciones entre sistemas adquieren cada vez mayor importancia, "*permitiendo a los repositorios ofrecer servicios de impacto a pesar de los pocos datos disponibles en los repositorios institucionales comparados con los sistemas globales*" (Tarver et al., 2014).



Señalar que la inclusión de identificadores de autor externos adicionales posibilitará una nueva generación de servicios orientados al investigador, identificándose diversas áreas de actuación (ORCID, 2014), entre las que destacaríamos de interés para el e-IEO las siguientes:

- Enlace desde el perfil del autor en e-IEO al registro ORCID en orcid.org
- Simplificación del autoarchivo, conectando este proceso con los registros de publicaciones de un autor a través de su identificador ORCID, evitando la doble introducción de datos y mejorando la calidad de los registros.
- Sincronización de publicaciones entre e-IEO y los registros orcid.org
- Uso del identificador ORCID como clave de autoridad de autor, cuando la adopción de este identificador sea generalizada por la comunidad investigadora.

VOCABULARIOS CONTROLADOS DE MATERIAS

Cuando los repositorios temáticos como el e-IEO se plantean ofrecer servicios y ser visibles por una audiencia más allá de sus fronteras institucionales, es esencial la descripción de contenidos y su indexado por medio de metadatos estandarizados que sean relevantes, tanto en el proceso de submisión como en el descubrimiento de recursos (Subirats *et al.*, 2012) y por esto es especialmente relevante el uso de tesauros temáticos, considerados éstos como un conjunto de “términos” empleados para representar los conceptos, temas o contenidos de los documentos de un disciplina o temática específica.

La incorporación de metadatos estandarizados por medio de este tipo de vocabularios controlados tiene las ventajas de a) mejorar las capacidades de búsqueda y descubrimiento de los ítems; b) mejorar o posibilitar la interoperabilidad (semántica) con repositorios o recolectores temáticos y c) incrementar la visibilidad de los ítems en repositorios y recolectores, por ejemplo en OCLC).

Los vocabularios controlados en dominios temáticos específicos tienen un amplio potencial cuando se usan en la recuperación de información, principalmente por su posibilidad de incorporar términos relacionados usados en el indexado, así como términos multi-lenguaje, (Borst, 2012) aspectos éstos que se han valorado en el repositorio e-IEO.

En línea con otros plataformas de repositorio de ciencias marinas, cuya referencia principal es AgriOcean DSpace, desarrollo conjunto de la FAO, agencia de las Naciones Unidas y de Unesco-IOC/IODE, se ha incorporado al e-IEO un subconjunto del vocabulario especializado AGROVOC. Igualmente se está evaluando la incorporación al repositorio en una fase posterior del vocabulario especializado Aquatic Sciences and Fisheries Abstracts (ASFA).



AGROVOC cubre todas las áreas de interés de la FAO, incluyendo alimentación, nutrición, agricultura, pesca, etc, por lo que la implementación para el e-IEO es un subconjunto del vocabulario, incorporando las subclasificaciones de recursos marinos, acuáticos y pesqueros, y que en forma de esquema SKOS-XL está disponible para organizaciones de todo el mundo.

Al ser un vocabulario que en su versión amplia maneja 32.000 conceptos en 21 idiomas, su tratamiento por DSpace tiene que realizarse, por motivos obvios de rendimiento, mediante la incorporación de un servidor de vocabularios específico, ASKOSI, mejorando significativamente la usabilidad de la solución construida.

REFERENCIAS BIBLIOGRÁFICAS

Borst, T. (2012). Usage and Impact of Controlled Vocabularies in a Subject Repository for Indexing and Retrieval. In *Liber Quarterly* Volume 21 Issue 3/4 2012 pp.: 445- 453.

DRIVER 2.0. (2008). Directrices para proveedores de contenido - Exposición de recursos textuales con el protocolo OAI-PMH. Digital Repository Infrastructure Vision for European Research.

http://www.driver-support.eu/documents/DRIVER_2_0_Guidelines_Spanish.pdf

ORCID (2014). ORCID Member Integration Guide.

<http://orcid.org/organizations/integrators>. Accessed July 30, 2014.

Subirats, I.; Malapela, T.; Dister, S.; Zeng, M.; Goovaerts, M.; Pesce, V. y Keizer, J. (2012). Reorienting open repositories to the challenges of the Semantic Web: Experiences from FAO's contribution to the resource processing and discovery cycle in repositories in the agricultural domain. In *Metadata and Semantics Research*, pp.: 158-167. Springer Berlin Heidelberg.

Tarver, Hannah; Waugh, Laura; Phillips, Mark Edward y Hicks, William. (2014). *Implementing Name Authority Control into Institutional Repositories: A Staged Approach*. UNT Digital Library.

<http://digital.library.unt.edu/ark:/67531/metadc172365/>. Accessed July 30, 2014.



Repositório de dados científicos

Grandes volumes de dados (Big Data), formatos de dados e metadados, reprodutibilidade e infraestrutura, aplicações para conjuntos de dados em acesso aberto



NECESIDADES DE GESTIÓN DE DATOS CIENTÍFICOS EN ARGENTINA. EL CASO DE LA UNIVERSIDAD NACIONAL DE ROSARIO¹

Paola Carolina Bongiovani¹, Luis Martínez-Uribe²

¹Facultad de Humanidades y Artes. Universidad Nacional de Rosario. Rosario, Argentina, LLAAR (Lista Latinoamericana de Acceso Abierto y Repositorios) pbongio@unr.edu.ar

²Fundación Juan March lmartinez@march.es

RESUMEN:

Recientemente se sancionó en Argentina Ley 26.899 "Repositorios digitales institucionales de Acceso Abierto" que establece la obligatoriedad de desarrollar repositorios digitales de acceso abierto, propios o compartidos, por parte de los organismos e instituciones públicas que componen el Sistema Nacional de Ciencia, Tecnología e Innovación. Esta ley requiere el establecimiento de políticas institucionales para la gestión, el acceso público y la preservación de datos primarios de investigación. En el marco de los proyectos "Hacia un modelo teórico, metodológico y tecnológico para el repositorio institucional de acceso abierto de la UNR fundamentado en las prácticas de su comunidad académica" (ING351) y "Hacia el desarrollo y utilización de Repositorios de Acceso Abierto para Objetos Digitales Educativos en el contexto de las universidades públicas de la región centro-este de Argentina" se realizó una encuesta a docentes investigadores de la Universidad Nacional de Rosario donde se relevó la existencia de producción de datos primarios de investigación, como así también las prácticas en relación a la utilización de éstos datos y las necesidades de servicios en relación a la gestión de los mismos. La presentación comunicará las principales experiencias internacionales en relación a gestión de datos primarios en universidades, el contexto nacional y los resultados obtenidos en la encuesta. Se expondrá sobre los tipos de datos científicos que se producen en la UNR, los formatos y paquetes de software utilizados por los investigadores para trabajar con datos de investigación, las modalidades utilizadas para compartir datos con sus colegas, los factores motivadores y desalentadores para compartir datos científicos en repositorios digitales y las necesidades de servicios de asesoramiento en relación a la gestión de datos. Asimismo se plantean propuestas de acciones a realizar en función de responder a las necesidades de los investigadores.

Palabras clave: Argentina, Universidad Nacional de Rosario, gestión de datos científicos, repositorios digitales



¹ Este trabajo se comparte bajo la licencia Creative Commons Atribución-NoComercial-CompartirIgual 2.0 Genérica (CC BY-NC-SA 2.0)



ABSTRACT:

Recently the Argentinean Senate passed Law 26.899 "Open Access Institutional Digital Repositories" requiring Open Access to publicly funded research outputs and establishing the obligation to develop open access digital repositories, own or shared by public agencies and institutions that make up the Science, Technology and Innovation National System. This legislation obliges the establishment of institutional policies for management, public access and preservation of scientific research data. As part of the project "Towards a theoretical , methodological and technological model for open access institutional repository at UNR based on practices of its academic community" (ING351) and "Towards the development and use of Open Access Repositories for Digital Educational objects in the context of public universities in the central- eastern Argentina " faculty researchers from the National University of Rosario were surveyed to uncover the existence of production of primary research data, as well as use practices of these data and service requirements relating to their management. The presentation will provide an overview of major international experiences regarding management of scientific data at universities, the national context and the results of the survey. The work covers the types of scientific data produced in the UNR , formats and software packages used by researchers to work with data, the methods used to share data with colleagues, motivations and disincentives to share data by scientists in digital repositories and needs for counseling regarding data management . It will discuss proposed actions to meet the needs of researchers.

Keywords: Argentina, National University of Rosario, scientific data management, digital repositories

INTRODUCCIÓN

La mayoría de los grandes retos científicos de nuestro tiempo necesitan de gran cantidad y diversidad de datos. De esta forma los datos primarios producidos por la investigación están cobrando una importancia y valor cada vez mayor. Compartir los datos de investigación brindando acceso a los mismos y permitiendo su reutilización genera muchas ventajas (Tenopir et al 2011):

- El análisis de los datos de investigación ayuda a verificar los resultados, que es un aspecto fundamental en el proceso de investigación;
- Diferentes interpretaciones o enfoques aplicados a datos existentes contribuyen a los avances científicos, especialmente en un contexto interdisciplinario;
- La preservación a largo plazo bien gestionada, permite mantener la integridad de los datos;
- Cuando se dispone de datos, la recolección de datos se reduce por lo que se optimiza el uso de recursos;



- La disponibilidad de datos ofrece garantías contra la mala conducta relacionada con la fabricación y falsificación de datos;
- La replicación de estudios con datos existentes sirve como herramienta para la formación de nuevos investigadores.

Un estudio internacional sobre prácticas y percepciones en relación al uso y compartición de datos de investigación que incluyó respuestas de 1329 científicos en todas las áreas temáticas (Tenopir et al, 2011) encontró que el 60% (promedio de todas las áreas) de los investigadores piensa que la falta de acceso a los datos generados por otros investigadores o instituciones es un obstáculo importante para el progreso de la ciencia. Los motivos por los cuales los datos de los investigadores estudiados no estaban disponibles por vía electrónica a los demás fue la falta de tiempo (54%), seguida de la falta de financiamiento (40%).

Gestión de datos primarios en universidades. Experiencias internacionales

Desde que en 2006 la OCDE lanzase su informe promoviendo nuevas prácticas de gestión de datos científicos, cada vez son más las agencias de financiación internacionales que exigen que los datos se gestionen y compartan. Por ejemplo, desde 2011 todas las propuestas de solicitud de financiamiento presentadas a la National Science Foundation (NSF) obligatoriamente deben contener un Plan de Gestión de Datos de investigación. Otras agencias de financiamiento de investigación en Estados Unidos como la National Aeronautics and Space Administration (NASA), National Oceanographic Data Center (NODC) y National Institutes of Health (NIH) contemplan la gestión y disponibilidad de datos resultantes de proyectos financiados por éstas.

Por otra parte, los Consejos de Investigación del Reino Unido también tienen políticas explícitas hacia los investigadores financiados en relación a la gestión, acceso y preservación de datos de investigación. Los planes de gestión de datos de investigación que exigen las agencias de financiamiento (DCC, 2013) contemplan aspectos comunes tales como: los tipos de datos que serán creados o recolectados (tipos, formatos, volumen de datos), el modo de creación o recolección de los datos (estándares, metodologías, versiones), metadatos y documentación que acompañarán los datos (estándares de metadatos, tecnología requerida para su uso), modalidad de gestión de cuestiones éticas (autorizaciones, protección de identidad, datos sensibles), aspectos relativos a la propiedad intelectual y derechos de autor de los datos, almacenamiento y backups de los datos, modo de acceso y seguridad de los datos, cómo se compartirán los datos y la existencia de restricciones, cuestiones relativas a la preservación de los datos, entre otros aspectos.



Estos requerimientos de las agencias de financiamiento alrededor del mundo exigen que las universidades generen nuevos servicios de datos de investigación para sus investigadores. La mayoría de las universidades y centros de investigación importantes alrededor del mundo cuenta con un repositorio institucional donde se difunde y preserva la producción científica (artículos, tesis, actas de congresos, etc.) de éstas instituciones. Según Hernández-Pérez y García-Moreno (2013) en un futuro próximo las instituciones también contarán con un repositorio de datos de investigación o tendrán que acordar con instituciones que cuenten con éstos. Pero los servicios de datos tienen que ir mas allá de los repositorios de datos que recopilen esos datos tras ser creados y explotados por los investigadores.

Martínez-Urbe y Macdonald (2008 & 2009) observan que las bibliotecas pueden jugar un rol importante en el apoyo a la gestión de datos generados y utilizados en la ciencia brindando servicios a lo largo del ciclo de vida de los datos. Los profesionales en este campo “data librarians” trabajan en conjunto con los investigadores ayudándolos en el tratamiento de datos, la elección de estándares y metadatos, así como con los aspectos legales y recomendando repositorios para depositar los datos científicos creados. Estos servicios de datos tendrán que poner en marcha apoyo para la búsqueda, análisis y visualización, como también estrategias de preservación, entre otros.

La creación de estos servicios de datos en instituciones científicas requiere de una coordinación estratégica y recursos a nivel nacional e internacional. Un buen ejemplo es Australia, en 2009 se estableció la iniciativa Australian National Data Service (ANDS) con un presupuesto de alrededor de 65 millones de euros para sus primeros seis años. Su objetivo es transformar el entorno de datos de investigación de Australia y permitir que la investigación con datos se reconozca como recurso estratégico a nivel nacional y se apoye a nivel institucional, nacional e internacional. Gracias a ANDS muchas universidades en Australia poseen servicios de datos en su mayor parte dirigidos desde las bibliotecas y en colaboración con los departamentos de Informática y de apoyo a la investigación. En Reino Unido y Estados Unidos agencias de financiación como JISC y NSF han destinado también cuantiosos fondos a programas de financiación dedicados a comprender mejor y desarrollar este tipo de servicios de datos que apoyan al investigador.

En 2012 la Asociación de Bibliotecas Universitarias y de Investigación (ACRL) realizó un relevamiento entre sus miembros en Estados Unidos y Canadá y encontraron que sólo una pequeña minoría de las bibliotecas universitarias en los Estados Unidos y Canadá ofrecen actualmente servicios de datos de investigación (RDS), pero un 25%-35% de todas las bibliotecas universitarias tienen previsto ofrecer algunos servicios dentro de los próximos dos años (Tenopir et al., 2012). La creación de guías web para ayudar a localizar datos es el servicio más comúnmente ofrecido o planeado. Las bibliotecas en instituciones grandes o en instituciones con investigadores que reciben fondos de la NSF son más proclives a ofrecer una gama de servicios más completos.



Las dificultades comunes encontradas en algunas instituciones al implementar servicios de gestión de datos (Raboin et al., 2012) son lograr apoyo institucional, gestionar la integración de los servicios con estructuras de personal existentes o nuevas, y lograr satisfacer las necesidades de los investigadores. Asimismo se plantea la necesidad de incorporar nuevos bibliotecarios al área de gestión de datos ya sea como nexo con los investigadores brindando servicios de consulta y para la gestión de metadatos.

Kouper et al.(2013) ofrecen una “hoja de ruta” para la implementación de servicios que soporten la creación, almacenamiento, distribución y preservación de datos de investigación cuyos componentes son Exploración del Entorno, Investigación de necesidades de Investigadores, Desarrollo de Herramientas, Políticas, Testeo, Evaluación y Valoración, Educación y Divulgación. Como primer paso de la hoja de ruta es la “Exploración del entorno” para lo cual realizaron entrevistas a representantes de universidades que cuentan con servicios de datos tales como Emory, Indiana, Purdue, Cornell, California Digital Library, Johns Hopkins, Virginia, Penn State, Illinois at Urbana-Champaign y Wisconsin-Madison. Los resultados de esta exploración indican que los servicios de datos deberían desarrollarse en torno a las investigaciones y los ciclos de vida de los datos, para lo cual se deben identificar recursos y servicios en cada momento de este ciclo de vida. Por otra parte se necesitan “agentes de cambio” que puedan apoyar y promover estos servicios de datos entre los investigadores y alumnos. La cooperación y colaboración aparecieron como componentes esenciales en el establecimiento de éstos servicios.

Iniciativas Nacionales

En Argentina existen políticas públicas en relación a la gestión y disponibilidad en abierto de los datos primarios de investigación. El Programa de Grandes Instrumentos y Bases de Datos² es una iniciativa conjunta del Ministerio de Ciencia, Tecnología e Innovación Productiva y el Consejo Interinstitucional de Ciencia y Tecnología (CiCyT) que tiene como objetivo la utilización eficiente de los grandes equipamientos y una mejor organización y acceso a las bases de datos científicos existentes en las distintas instituciones de ciencia y tecnología de todo el país. Los sistemas de datos existentes son el Sistema Nacional de Datos Biológicos, el Sistema Nacional de Datos del Mar, el Sistema Nacional de Datos Climáticos y el Sistema Nacional de Repositorios Digitales.

Recientemente se sancionó una ley de repositorios digitales que establece la obligatoriedad de desarrollar repositorios digitales de acceso abierto, propios o compartidos, por parte de los organismos e instituciones públicas que componen el Sistema Nacional de Ciencia, Tecnología e Innovación (Bongiovani y Nakano, 2011; Honorable Cámara de senadores de la Nación, 2013³). La mencionada ley requiere el establecimiento de políticas institucionales para la gestión, el acceso público y la preservación de datos primarios de investigación.

² <http://sistemasnacionales.mincyt.gob.ar/>

³ <http://www.senado.gov.ar/parlamentario/comisiones/verExp/26.12/CD/PL>



El Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) viene desarrollando el proyecto Plataforma Interactiva de Investigación en Ciencias Sociales (PLIICS) a fin de facilitar el acceso a mayor cantidad de información primaria en Ciencias Sociales y Humanidades y su interrelación de modo de fortalecer el enfoque multidisciplinarios y la comprensión de problemas complejos.

Los resultados de la encuesta sobre datos de la PLIICS (CONICET, 2011⁴, Pluss y Leff, 2013) muestran que la gran mayoría de investigadores producen datos en formato digital con una gran variedad de aplicaciones de software y de formatos. La mitad de los encuestados está de acuerdo en que estos datos se compartan a partir de los dos primeros años o antes. Sin embargo pocos de los datos llegan a enriquecerse con metadatos y pocos de los investigadores conocen de la existencia de repositorios digitales en sus lugares de trabajo.

La Universidad Nacional de Rosario (UNR), que se encuentra entre las principales universidades nacionales de Argentina en cuanto a producción científica y académica, cuenta con su RI, denominado RepHipUNR que permite archivar, organizar, preservar y distribuir digitalmente en variados formatos tanto materiales de enseñanza y aprendizaje como la producción científica de Investigación y Desarrollo (I+D) de los profesores, profesionales e investigadores de la UNR (Bongiovani, 2013). El RepHipUNR brinda acceso a artículos de revistas, documentos de conferencias, libros, capítulos de libros, tesis, materiales educativos, entre otros. Recientemente el repositorio adhirió al Sistema Nacional de Repositorios Digitales (SNRD). Asimismo la Facultad de Ciencias Agrarias y la Facultad de Facultad de Ciencias Bioquímicas y Farmacéuticas de la UNR han adherido al Sistema Nacional de Datos Biológicos.

OBJETIVO

Teniendo en cuenta la existencia de una ley que exige la creación de una política para la gestión de datos a nivel nacional, el presente trabajo tiene como objetivo identificar la existencia de producción de datos primarios de investigación y las prácticas en relación a los datos de investigación de los investigadores de la Universidad Nacional de Rosario y las necesidades de servicios en relación a la gestión de los mismos. En el marco de los proyectos "Hacia un modelo teórico, metodológico y tecnológico para el repositorio institucional de acceso abierto de la UNR fundamentado en las prácticas de su comunidad académica" (ING351UNR-CLACSO) y "Hacia el desarrollo y utilización de Repositorios de Acceso Abierto para Objetos Digitales Educativos en el contexto de las universidades públicas de la región centro-este de Argentina" se realizó una encuesta a docentes investigadores de la UNR.

⁴ <http://web.conicet.gov.ar/web/conicet.trabajar.investigacion/programa-ciencias-sociales>



MATERIALES Y METODOLOGÍA

El estudio se basa en las respuestas de 783 docentes investigadores de la Universidad Nacional de Rosario que participaron de la encuesta online “Acceso Abierto a las producciones académicas. Necesidades y Prácticas” implementada a través de SurveyMonkey (es.surveymonkey.com) entre el 5 de julio y el 31 de octubre de 2012 en el marco de los proyectos mencionados. La encuesta constó de 28 preguntas de las cuales se seleccionaron para este estudio las específicas en el apartado Datos Primarios:

Preguntas	Valores de las Respuestas
En el curso de su investigación, ¿genera datos primarios? (#18)	SI /NO
Los datos producidos, ¿se encuentran en formato digital? (#19)	Sí - Todos / Sí - Parcialmente / NO
¿Qué tipo de datos produce? Por favor seleccione todo lo que se aplique: (#20)	Datos experimentales / Observaciones / Simulaciones / Audio / Texto / Bases de datos estadísticos / Bases de datos documentales / Cartografía temática / Datos radiográficos / Datos generados por Sensores remotos / Datos estadísticos / Datos sintéticos / Datos instrumentales / Datos epidemiológicos / Datos topográficos / Dibujos / Entrevistas / Datos geofísicos / Fotografías / Historias de vida / Imágenes / Objetos (arqueológicos, antropológicos, etc.) / Planos, Mapas / Resultados de Cuestionarios cuantitativos / Resultados de Cuestionarios cualitativos / Registros de campo / Video / Otro
¿Qué formatos digitales emplea para crear, archivar y/o resguardar los datos primarios? Por favor seleccione todas las opciones que se apliquen: (#21)	Archivos de Bases de Datos (Ej. Access, MySQL) / CAD/GIS/ Extensible mark-up language (XML)/ Archivos planos (Ej. FITS)/ Hypertext mark-up language (HTML)/ Texto plano (.txt)/ Portable document format (.pdf) / Rich text files (.rtf) / Planilla de cálculo (Ej. Excel) / Software estadístico (Ej. SPSS) / Tablas / Archivo de procesamiento de texto (Ej. Word .doc) / Formatos de imagen (.bmp.; .jpg; .tiff, etc)/ Formatos de imagen vectorial (.cdr; .ai; etc.) / Formatos de audio. (Mp3/ wav/ etc.)/ Formatos en video. (avi; .mpeg, etc.) / Otro
¿Qué aplicaciones de software utiliza para trabajar con sus datos de investigación? Por favor seleccione todas las opciones que se apliquen (#22)	Procesadores de texto / Hojas de cálculo / Software de bases de datos / Software para diseño gráfico / Software para transcripción de audio y video / Software de edición de audio y video / Software de edición de imágenes / Software de tratamiento de datos cualitativos / Software de tratamiento de datos cuantitativos / Otro



Actualmente, ¿de qué manera otros investigadores pueden acceder a sus datos? Por favor seleccione todas las opciones que se apliquen (#23)	En formato impreso / Datos distribuidos vía email / Intercambio de dispositivos de almacenamiento portátiles (CD, USB, etc.) / Subo mis datos a un servidor web y envío un password a mis colegas / Subo mis datos a un servidor web y publicito el URL de acceso / A través del editor de una revista / No pueden acceder a mis datos / Otro
¿Qué factores lo motivarían para que comparta sus datos primarios a través de un repositorio digital? Por favor seleccione todas las opciones que se apliquen (#24)	Requerimiento por parte de quienes financian mi investigación / Beneficios demostrables para mi carrera (mejora en mi evaluación, obtener financiamiento futuro, obtener ideas de investigación) / Beneficios demostrables para mi institución / Beneficios potenciales para la comunidad de investigación / Beneficios potenciales para la sociedad / Permitir colaboraciones y contribuciones de otros / Mayor validación de los resultados de investigación / Mayor visibilidad de mi investigación / Otro
¿Qué factores lo desalentarían de compartir sus datos primarios a través de un repositorio digital? Por favor seleccione todas las opciones que se apliquen (#25)	Falta de financiamiento para la tarea de estandarizar mis datos para su acceso / Tiempo y esfuerzo requerido para compartir mis datos / Desconocimiento de estándares para compartir mis datos / Restricciones éticas / Restricciones relativas a la seguridad y confidencialidad de los datos / Otro
¿Estaría usted interesado en contar con asesoramiento sobre la gestión de sus datos? Por favor seleccione todas las opciones que se apliquen (#26)	Asesoramiento sobre cuestiones prácticas relacionadas con la gestión de datos / Asesoramiento sobre digitalización, herramientas y servicios / Asesoramiento sobre la creación de un plan de gestión de datos de investigación / Asesoramiento sobre opciones para almacenar, gestionar y compartir datos de forma segura / Asesoramiento sobre preservación de mis datos de investigación / Otros

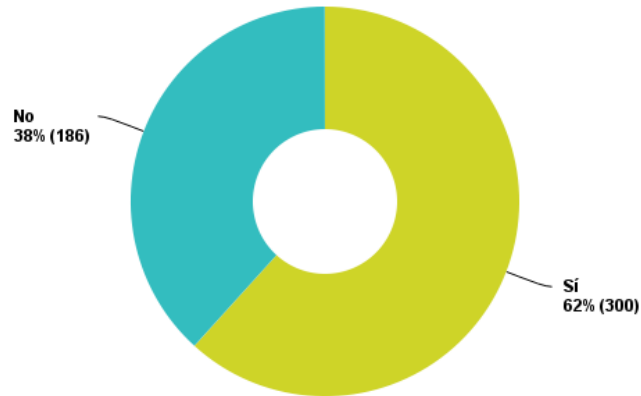
RESULTADOS

Producción de datos de investigación y características

Un 62% de los investigadores que contestaron la pregunta dice que produce datos en el curso de su investigación y de éstos un 87% tiene los datos en formato digital total (34%) o parcialmente (52%).

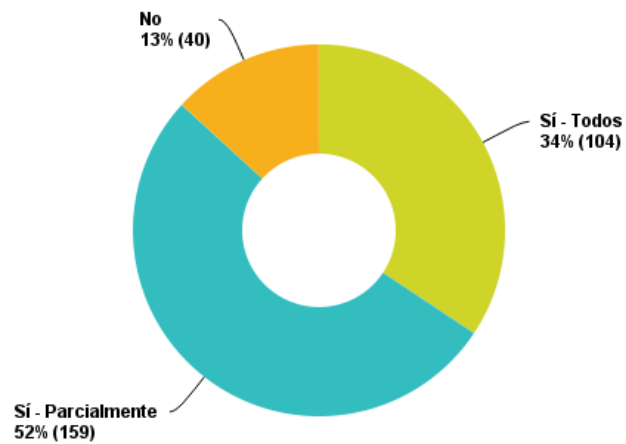
**Q18 En el curso de su investigación,
 ¿genera datos primarios?**

Respondido: 486 Omitido: 297



**Q19 Los datos producidos, ¿se encuentran
 en formato digital?**

Respondido: 303 Omitido: 480



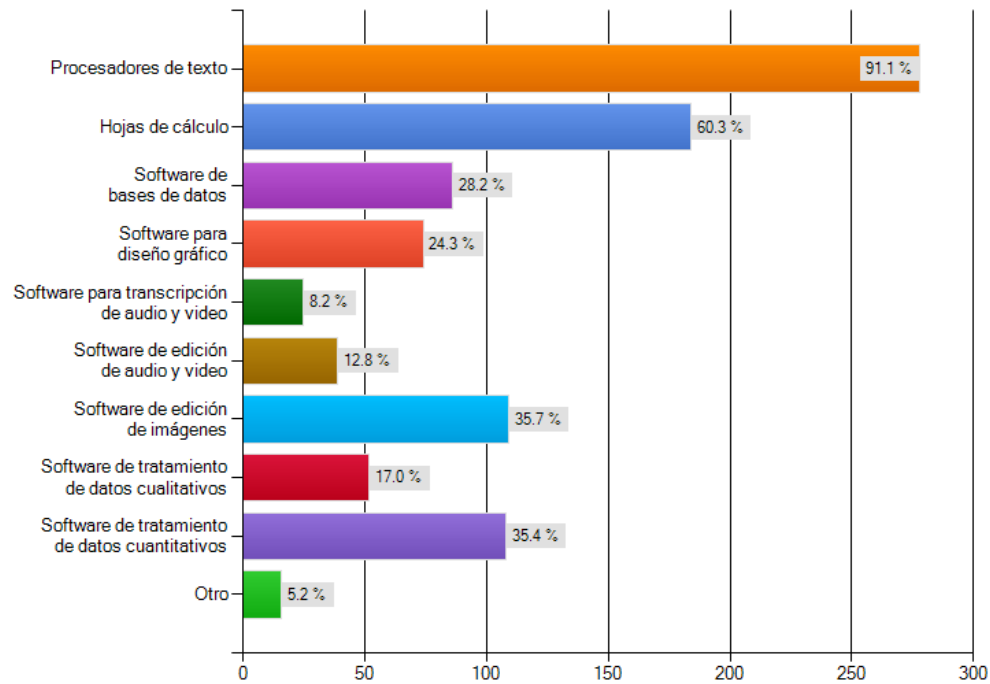
Entre los tipos de datos que producen los investigadores se encuentran datos experimentales (46%), observaciones (45%), texto (41%), datos estadísticos (31%) entre otros muchos tipos identificados en la siguiente tabla:

Opciones de respuesta	Respuestas	
Datos experimentales	46%	140
Observaciones	45%	137
Texto	41%	126
Datos estadísticos	31%	94
Imágenes	27%	82
Entrevistas	25%	77
Fotografías	22%	68
Resultados de Cuestionarios cualitativos	22%	68
Registros de campo	20%	60
Bases de datos estadísticos	18%	55
Resultados de Cuestionarios cuantitativos	15%	46
Bases de datos documentales	14%	42
Vídeo	11%	35
Simulaciones	10%	32
Datos epidemiológicos	10%	30
Historias de vida	10%	30
Audio	9%	28
Datos instrumentales	6%	19
Dibujos	5%	14
Datos radiográficos	3%	10
Planos, Mapas	3%	9
Otro	2%	6
Objetos (arqueológicos, antropológicos, etc.)	2%	5
Datos generados por Sensores remotos	1%	4
Datos sintéticos	1%	4
Cartografía temática	1%	3
Datos topográficos	1%	2
Datos geofísicos	0%	1
Total de encuestados: 306		

Los formatos más usados por éstos investigadores para crear, archivar y/o resguardar los datos generados son archivos de procesamiento de texto (75%), planilla de cálculo (55%), PDF (49%), formatos de imagen y tablas (48%), entre otros formatos. Las aplicaciones de software más utilizadas para trabajar con los datos de investigación son procesadores de texto (91%), hojas de cálculo (60%), software de edición de imágenes (36%) y software de tratamiento de datos cuantitativos (35%) entre otros.

Opciones de respuesta	Respuestas	
Archivo de procesamiento de texto (Ej. Word/.doc)	75%	227
Planilla de cálculo (Ej. Excel)	55%	166
Portable document format (.pdf)	49%	149
Formatos de imagen (.bmp.; .jpg; .tiff, etc)	49%	147
Tablas	30%	90
Texto plano (.txt)	22%	68
Software estadístico (Ej. SPSS)	20%	60
Formatos de audio. (Mp3, wav, etc.)	17%	51
Formatos en video. (avi; .mpeg, etc.)	15%	44
Archivos de Bases de Datos (Ej. Access, MySQL)	14%	41
Rich text files (.rtf)	7%	22
Formatos de imagen vectorial (.cdr; .ai; etc.)	6%	19
Hypertext mark-up language (HTML)	5%	16
Otro	5%	16
CAD/GIS	2%	7
Extensible mark-up language (XML)	2%	5
Archivos planos (Ej. FITS)	1%	4
Total de encuestados: 303		

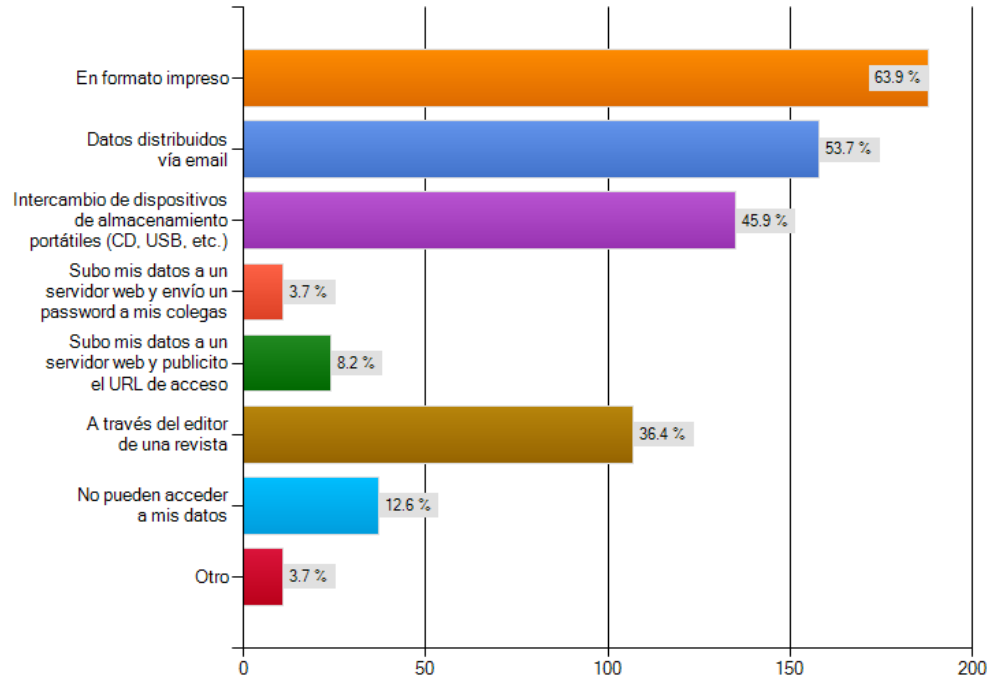
¿Qué aplicaciones de software utiliza para trabajar con sus datos de investigación? Por favor seleccione todas las opciones que se apliquen:



Modalidad y motivaciones para compartir los datos de investigación

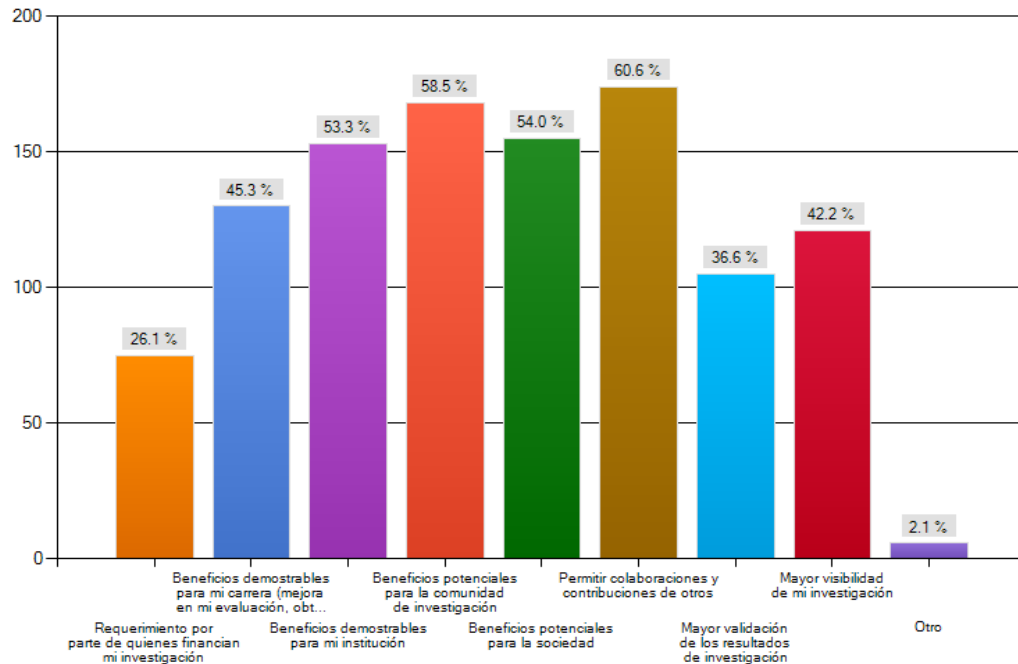
El formato impreso es el modo más común para compartir los datos de investigación que generan con otros investigadores (64%), también se comparte vía correo electrónico (54%) , a través de dispositivos electrónicos portátiles (46%) y a través del editor de una revistas (36%). Un 13% dice que otros investigadores no pueden acceder a sus datos.

Actualmente, ¿de qué manera otros investigadores pueden acceder a sus datos? Por favor seleccione todas las opciones que se apliquen:



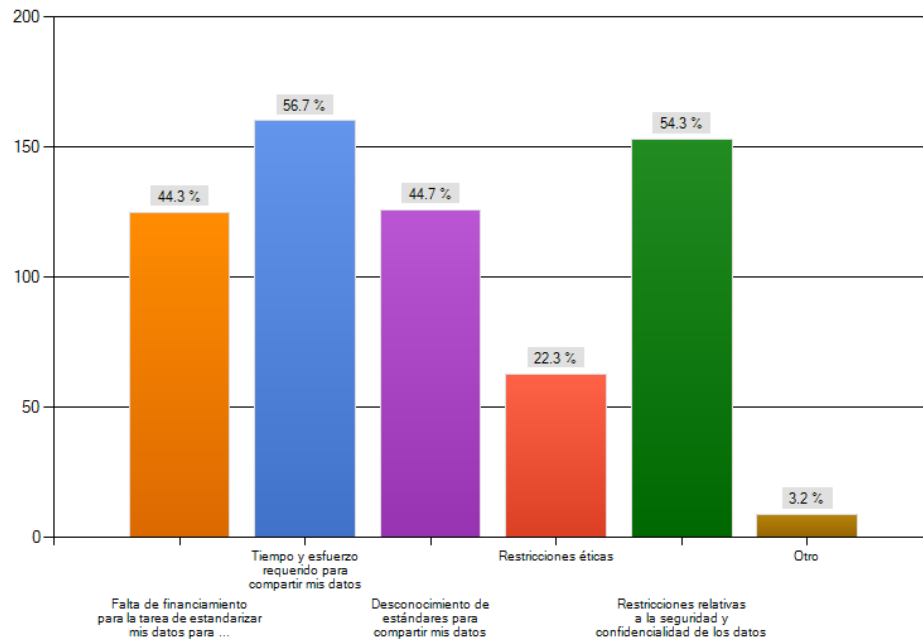
Entre los factores que más motivarían a los investigadores a compartir sus datos de investigación a través de repositorios digitales, se cuenta permitir colaboraciones y contribuciones de otros (61%), beneficios potenciales para la comunidad de investigación (58%), beneficios potenciales para la sociedad (54%), beneficios demostrables para la institución de pertenencia (53%) y beneficios demostrables para la carrera del investigador (45%).

¿Qué factores lo motivarían para que comparta sus datos primarios a través de un repositorio digital? Por favor seleccione todas las opciones que se apliquen:



El tiempo y esfuerzo requerido para compartir los datos es el principal factor que desalienta a los investigadores de la UNR (57%), así como restricciones relativas a la seguridad y confidencialidad de los datos (54%), desconocimiento de estándares para compartir sus datos (45%) y falta de financiamiento para la tarea de estandarizar mis datos para su acceso (44%).

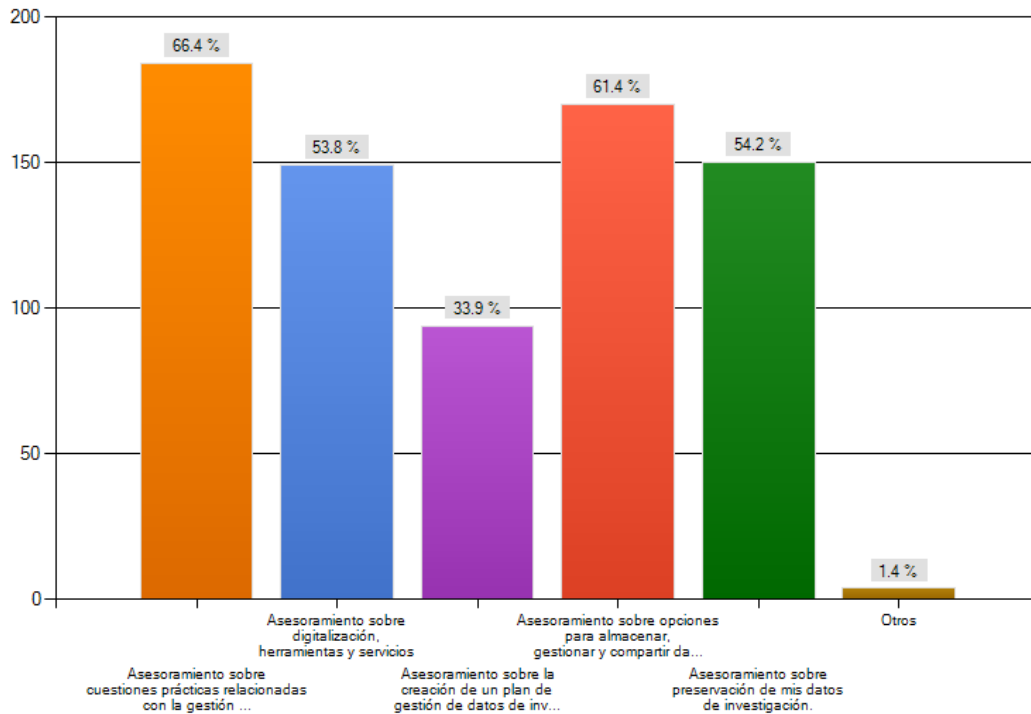
¿Qué factores lo desalentarian de compartir sus datos primarios a través de un repositorio digital? Por favor seleccione todas las opciones que se apliquen:



Interés en servicios para la gestión de datos

Los investigadores que contestaron la encuesta indicaron que están interesados en contar con asesoramiento sobre cuestiones prácticas relacionadas con la gestión de datos (66%), sobre opciones para almacenar, gestionar y compartir datos de forma segura (61%), sobre preservación de datos de investigación (54%) y sobre digitalización, herramientas y servicios (54%) entre otros.

¿Estaría usted interesado en contar con asesoramiento sobre la gestión de sus datos? Por favor seleccione todas las opciones que se apliquen:



DISCUSIÓN

En Argentina, el Ministerio de Ciencia, Tecnología e Innovación Productiva, a través del Programa de Grandes Instrumentos y Bases de Datos⁵ ofrece financiamiento y una estructura que permite la coordinación nacional de las instituciones que generan datos de investigación en áreas temáticas específicas donde la UNR ya participa en diferentes áreas temáticas.

Asimismo la aprobación de la Ley 26.899 “Repositorios digitales institucionales de Acceso Abierto” donde se requiere el establecimiento de políticas institucionales para la gestión, el acceso público y la preservación de datos primarios de investigación, obliga a las universidades a desarrollar acciones en esta línea de modo de cumplir con la ley.

Por su parte, el relevamiento realizado da cuenta de los datos con los que están trabajando los investigadores de la Universidad Nacional de Rosario y de las necesidades planteadas para su gestión.

⁵ <http://sistemasnacionales.mincyt.gob.ar/>



Se debería trabajar en el establecimiento de un Comité de Trabajo en el que se involucre en las decisiones a todas las áreas pertinentes y personal con experticia en la materia.

Se propone convocar a un número reducido proyectos de investigación que tengan un alto interés en avanzar en la gestión de sus datos y que sirvan como proyectos piloto.

De este modo se podrá empezar a conocer mejor las características y los ciclos de vida de colecciones de datos específicas y empezar a esbozar políticas adecuadas para su gestión. Por su parte podrán surgir las necesidades más específicas de infraestructura de repositorios de datos adaptada a las necesidades de la institución.

Es importante destacar que estas nuevas infraestructuras y servicios requerirán financiamiento de la propia institución y también recursos externos a través del Ministerio de Ciencia, Tecnología e Innovación Productiva para sostenerse en el tiempo. Es necesario invertir en capacitación de profesionales en la gestión de datos (data librarians), en la puesta en marcha de servicios de consulta y asesoramiento a los investigadores, además de la infraestructura necesaria para almacenar, dar acceso y preservar esos datos de investigación.

Por otra parte, se deberá trabajar en la sensibilización de los investigadores de modo de generar un cambio cultural, mostrando las ventajas de compartir sus datos de investigación.

AGRADECIMIENTOS

A los equipos de investigación de los proyectos del “Hacia un modelo teórico, metodológico y tecnológico para el repositorio institucional de acceso abierto de la UNR fundamentado en las prácticas de su comunidad académica” (ING351UNR-CLACSO) y “Hacia el desarrollo y utilización de Repositorios de Acceso Abierto para Objetos Digitales Educativos en el contexto de las universidades públicas de la región centro-este de Argentina” (PICTO-CIN N° 0143-ANPCyT) ambos parte del Programa Dispositivos Hipermediales Dinámicos por los aportes en la evaluación del instrumento de la encuesta y su puesta en marcha.

BIBLIOGRAFÍA

Bongiovani, Paola y Nakano, Silvia (2011) Acceso abierto en Argentina: la experiencia de articulación y coordinación institucional de los repositorios digitales en ciencia y tecnología. E-colabora, vol. 1, no 2, p. 163-179. Disponible en <http://publicaciones.renata.edu.co/index.php/RCEC/article/view/56>

DCC. (2013). Checklist for a Data Management Plan. v.4.0. Edinburgh: Digital Curation Centre. Available online: <http://www.dcc.ac.uk/resources/data-management-plans>



Hernández-Pérez, T., & García-Moreno, M. A. (2013). Datos abiertos y repositorios de datos: nuevo reto para los bibliotecarios. *El profesional de la información*, 22(3), 259-263.

Kouper, I., Akers, K. G., Nicholls, N. H., & Sferdean, F. C. (2013). A roadmap for data services. In *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries* (pp. 375-376). ACM.

Martínez-Uribe, L., & Macdonald, S. (2008). Un nuevo cometido para los bibliotecarios académicos: data curation. *El profesional de la información*, 17(3), 273-280.

Martinez-Uribe, L., & Macdonald, S. (2009). User engagement in research data curation. In *Research and advanced technology for digital libraries* (pp. 309-314). Springer Berlin Heidelberg.

Pluss, R. Y Leff, L. (2013). Plataforma Interactiva de Investigación para las Ciencias Sociales42 JAIIO Jornadas Argentinas de Informática.

Steinhart, G., Chen, E., Arguillas, F., Dietrich, D., & Kramer, S. (2012). Prepared to Plan? A Snapshot of Researcher Readiness to Address Data Management Planning Requirements. *Journal of eScience Librarianship*, 1(2), 1.

Raboin, R., Reznik-Zellen, R. C., & Salo, D. (2013). Forging New Service Paths: Institutional Approaches to Providing Research Data Management Services. *Journal of eScience Librarianship*, 1(3), 2.

Soehner, C., Steeves, C., & Ward, J. (2010). E-Science and Data Support Services: A Study of ARL Member Institutions. *Association of Research Libraries*.

Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu L, et al. (2011) Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE* 6(6): e21101. doi:10.1371/journal.pone.0021101

Tenopir, C., Birch, B., & Allard, S. (2012). *Academic Libraries and Research Data Services: Current Practices and Plans for the Future; an ACRL White Paper*. Association of College and Research Libraries, a division of the American Library Association.

Tenopir, C., Sandusky, R. J., Allard, S., & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*.

Wilson, J. A., Martinez-Uribe, L., Fraser, M. A., & Jeffreys, P. (2011). An Institutional Approach to Developing Research Data Management Infrastructure. *International Journal of Digital Curation*, 6(2), 274-287.



VISIBILIDADE DE DOCUMENTOS EM ACESSO ABERTO: NECESSIDADE DE EDUCAÇÃO DO USUÁRIO

Jorge Santa Anna

RESUMO

Discute aspectos relacionados à disponibilização de documentos digitais e sua utilização pelos usuários em espaço aberto. Para tanto, contextualiza a gênese e desenvolvimento do ciberespaço na sociedade moderna. Apresenta o contexto de desenvolvimento e as funcionalidades do acesso aberto. Analisa o grau de conhecimento dos usuários a respeito do uso das tecnologias digitais e o acesso a documentos livres, de forma lícita. Expõe a importância e necessidade de adequações dos ambientes digitais, face aos obstáculos enfrentados pelos usuários, bem como a capacitação deles na utilização de documentos disponibilizados em espaço aberto. Metodologicamente, foram utilizados procedimentos de pesquisa bibliográfica, documental e estudo em campo, através de entrevistas a um grupo de pesquisadores e profissionais. Por fim, concluiu-se que é preciso estabelecer políticas de acesso a esses ambientes, bem como a divulgação de como utilizar os documentos sem impactar com problemas de direitos autorais. Também se infere a necessidade de capacitar o usuário da informação eletrônica, tendo em vista sua adequação aos limites impostos pelo ambiente virtual.

Palavras-chaves: Novas tecnologias. Acesso aberto. Direitos autorais. Políticas de acesso. Capacitação.



ABSTRACT

Discusses aspects related to the availability of digital documents and its use by users in open space. To do so, contextualizes the genesis and development of the cyberspace in modern society. Presents the context of development and the functionalities of open access. Analyzes the degree of users' knowledge about the use of digital technologies and access to free documents, lawfully. Exposes the importance and need of adjustments of digital environments, due to the obstacles faced by users, as well as enabling them in the use of documents available in open space. Methodologically, were used procedures of bibliographical and documentary research, and field study by interviewing a group of researchers and professionals. At last, it was concluded that we need to establish policies of access to these environments, as well as the disclosure of how to use the documents without impacting with copyright issues. Also infers the need to empower the user of electronic information, in view of their adequacy to limits imposed by the virtual environment.

Keywords: New technologies. Open access. Copyright. Access policies. Training.

INTRODUÇÃO

O desenvolvimento tecnológico e os paradigmas instituídos pela sociedade contemporânea favorecem a concepção de ideias inéditas a respeito da disseminação da informação e seu uso pela sociedade. Essa ocorrência é fruto das novas necessidades impostas pelo espaço social, pautado na informação como insumo básico no desenvolvimento pessoal, cultural e econômico de uma nação.

Essa realidade viabiliza o fortalecimento de práticas em prol do acesso democrático à imensa onda de informações que permeiam o atual contexto. Assim, inúmeras questões são levantadas a respeito do movimento do acesso aberto, sobretudo, com o desenvolvimento da internet, que favoreceu o surgimento de um espaço interativo de acesso e compartilhamento de informações, tornando mais fácil o processo de comunicação científica.

Nesse contexto, tendo como base a pesquisa de Weitzel (2006), entende-se que esse movimento cria estruturas viáveis que permitem o acesso irrestrito à produção científica legítima, alterando não somente o processo de aquisição de informação científica, mas também a sua produção, disseminação e uso.

O movimento do acesso aberto possui uma proposta favorável à democratização do conhecimento na sociedade, garantindo, de modo geral, a concretização da cidadania, *efetivando* o cumprimento de direitos e deveres. No entanto, muitos problemas vêm à tona sobre a disponibilização de informações eletrônicas, problemas esses relacionados, sobremaneira, a uma nova conjuntura refletida na indústria editorial.

Está evidente que a consolidação do acesso aberto impacta nos processos e fluxos que permeiam a comunicação científica, favorecendo a consolidação de novas ambiências. Assim, a informação passa a ser transferida por meio de diferentes tipos de canais, envolvendo fluxos variados, por conseguinte, redefinindo novos papéis atribuídos a todos os profissionais que atuam na produção científica.

O estudo de Mueller (2006) enfatiza que o movimento para o acesso aberto ao conhecimento científico pode ser considerado como o fato mais interessante e talvez importante de nossa época no que se refere à comunicação científica. Ao mesmo tempo, este movimento representa enorme desafio para a comunidade científica, à medida que, quanto mais amplo o seu sucesso, mais radical será a mudança provocada no sistema tradicional e profundamente arraigado de comunicação do conhecimento científico.

No entendimento de Johnson (2002 apud LEITE, 2009), na conjuntura atual, a indústria editorial está sendo redefinida. Isso porque cresce claramente o papel de modelos alternativos de comunicação científica, tais como repositórios institucionais e bases de dados, ao quebrarem monopólios de editores e aumentarem a atenção e ciência de pesquisadores sobre a produção intelectual das universidades e institutos de pesquisa.

Sendo assim, este estudo atrela-se à temática do acesso aberto, enfatizando questões relacionadas à ampliação da sua visibilidade, garantindo a legitimidade dessas práticas. A questão norteadora na qual se apoia o referido estudo visa a encontrar respostas a respeito das práticas e percepções de usuários envolvidos no processo de comunicação científica quanto aos desafios impostos pelo ciberespaço.

Por decorrência, este artigo objetiva discutir aspectos relacionados à disponibilização de documentos digitais e sua utilização pelos usuários em espaço aberto. Para tanto, o estudo visa a contextualizar a gênese e desenvolvimento do ciberespaço na sociedade moderna; apresentar o contexto de desenvolvimento e as funcionalidades do acesso aberto; analisar o grau de conhecimento dos usuários a respeito do uso das tecnologias digitais e o acesso a documentos livres, de forma lícita. Além disso, expor a importância e necessidade de adequações dos ambientes digitais, face aos obstáculos enfrentados pelos usuários, bem como à sua capacitação para uso de documentos disponibilizados em espaço aberto.

Com base nessa realidade, o estudo torna-se pertinente, uma vez que a informação disponibilizada em novos artefatos tecnológicos desencadeia o surgimento de instrumentos até então não utilizados e adequação de técnicas e metodologias, assim como provoca também a necessidade de controle das condutas que precisam ser gerenciadas a fim de garantir o uso legal da informação eletrônica.

Metodologicamente, foram utilizados procedimentos de pesquisa bibliográfica, documental e estudo de campo, através de entrevistas a um grupo de pesquisadores, com vistas a coletar dados a respeito das práticas e percepções desses sujeitos no contexto do acesso aberto.



O estudo bibliográfico foi sustentado por autores renomados da área da Ciência da Informação, utilizando-se como materiais de pesquisa, livros e artigos por eles publicados. Para a análise documental, utilizaram-se fontes jurídicas, como Constituição Federal e a Lei de Direitos Autorais (Lei 9610). Já o estudo de campo teve como espaço de investigação, uma universidade federal, tendo como sujeitos de pesquisa, indivíduos vinculados a essa instituição: três pesquisadores, um profissional bibliotecário e um editor de revista eletrônica.

A técnica para coleta de dados foi a entrevista aplicada a esses sujeitos, de forma individualizada, sendo as informações fornecidas devidamente anotadas e gravadas, a fim de facilitar o processo de apresentação e interpretação dos dados coletados.

DIÁLOGO COM ALGUNS TEÓRICOS

A utilização da *internet* desencadeou inúmeras transformações que afetaram a sociedade em todas as suas dimensões, interferindo no comportamento de indivíduos e instituições. No âmbito da indústria da informação, a utilização de novos canais de transferência desencadeou diferentes formas de compartilhamento, proporcionando novas ambiências no processo de comunicação científica.

Se antes, as pesquisas eram compartilhadas através de processos tradicionais, como o uso de cartas, acarretando morosidade na produção de conhecimentos, na atualidade, o uso dos recursos fornecidos pela *internet* torna esse processo cada dia mais eficaz, podendo a informação ser socializada sem considerar barreiras geográficas (MUELLER, 2000).

A expansão da *internet* provocou mudanças comportamentais, havendo a necessidade de redefinição de muitos processos de trabalho, caracterizando a sociedade atual como um espaço integrado em redes, em que as informações são compartilhadas instantaneamente e em várias localidades (CASTELLS, 2007). Desse modo, desperta-se a formação de um novo espaço social, o ciberespaço, caracterizado como

[...] um novo espaço de interação humana que já tem uma importância enorme, sobretudo no plano econômico e científico e, certamente, essa importância vai ampliar-se e vai estender-se a vários outros campos [...]. O espaço cibernético é a instauração de uma rede de todas as memórias informatizadas e de todos os computadores (LÉVY, 1999, p. 29).

Nessa nova realidade, propaga-se a Teoria do Rizoma, ao defender que as instituições pertencentes à sociedade contemporânea, inseridas em um contexto de globalização, tendo auxílio das novas tecnologias, estão atreladas umas às outras, compartilhando informações em um processo contínuo, sem delimitação de início e fim (DELEUZE; GUATARY, 1995).

O aumento da produção bibliográfica foi condicionado, em grande parte, pelo avanço das novas tecnologias e com o uso da *internet*, os quais viabilizaram a criação de novas fontes de informação, bem como a transmissão da informação. De modo especial, os periódicos científicos têm sido uma das fontes que mais cresceu nas últimas décadas, representando um importante meio de comunicação a ser utilizado por autores, publicadores, bibliotecários e usuários (PACKER, 2005).

O mesmo autor defende que, a disponibilização do periódico em espaço digital desencadeou as primeiras resistências ligadas à defesa do papel como suporte de registro e publicação, que foram paulatinamente superadas devido às extraordinárias facilidades, potencialidades e conveniências oferecidas pelo acesso *online* aos artigos e outros textos.

Consolidam-se, na contemporaneidade, novos desafios, tanto para pesquisadores, autores e consumidores de informação, quanto para os profissionais e as instituições que atuam no controle, armazenamento e disponibilização das fontes eletrônicas.

Especificamente no contexto das bibliotecas e instituições de pesquisa, essas unidades vêm se constituindo como um espaço destinado a fornecer o acesso a todo tipo de fonte de informação a qualquer usuário, desconsiderando as limitações de tempo e espaço. Esse processo de interação dessas unidades com seus públicos, sem medir as fronteiras geográficas, somente é conseguido com o apoio das novas tecnologias da informação e comunicação (TIC) (SANTA ANNA, 2013).

Vê-se que as novas formas de disponibilização da informação, favorecendo diferentes mecanismos de comunicação, constituem um dos fatores mais interferentes na indústria da informação. Isso provoca, simultaneamente, inúmeras vantagens, sobretudo no que se refere à ampliação do acesso, como também pode representar obstáculos, face às novas estruturas tecnológicas e as formas e condições de seu uso.

Dentre os vários fatores que interferem no processo de comunicação científica, o movimento para o acesso aberto ao conhecimento científico vem sendo estudado com grande frequência, devido a sua importância (MUELLER, 2006). Segundo Oliveira e Gomes (2013), esse movimento é constituído por um conjunto de práticas diferenciadas da tradicional maneira de comunicar o conhecimento científico, no que se refere ao acesso às publicações periódicas condicionadas ao pagamento de caras assinaturas. O direito autoral tem se mostrado um dos principais entraves para a expansão do movimento de acesso aberto.

Assim, o acesso aberto, nesse contexto, significa

[...] a disponibilização livre pública na *Internet*, de forma a permitir a qualquer usuário a leitura, download, cópia, distribuição, impressão, busca ou criação de links para os textos completos dos artigos, bem como capturá-los para indexação ou utilizá-los para qualquer outro propósito legal [...] (LEITE, 2009, p. 15, grifo nosso).



Mueller (2006, p. 27) esclarece que, embora a proposta do acesso aberto vise a consolidar a disseminação do conhecimento, viabilizando sua produção, face às facilidades de compartilhamento, ele provoca mudanças impactantes. Segundo a mesma autora, a chegada do periódico científico eletrônico, na década de 90, despertou “[...] esperanças, em muitos pesquisadores, de uma mudança radical no sistema tradicional de comunicação científica”.

No decorrer dos tempos, o movimento do acesso aberto preocupa pesquisadores do mundo todo, sendo motivo de discussão entre especialistas envolvidos no mercado editorial, além de ser objeto de discussão em eventos de diferentes áreas. Nos anos de 2000 a 2003 foram realizados, nas cidades de Budapeste, Bethesda e Berlim, três importantes eventos em prol do acesso aberto às publicações científicas (OLIVEIRA; GOMES, 2013).

De modo geral, esses eventos resultaram na elaboração de três declarações que, posteriormente, constituiriam relevantes manifestações em favor do acesso aberto às publicações científicas no mundo. Os referidos eventos convocaram autores, associações científicas, editores, universidades e bibliotecas a trabalharem na remoção das barreiras que impedem o acesso livre à literatura científica (OLIVEIRA; GOMES, 2013).

Como resultado da Declaração de Budapeste, sendo consolidado, posteriormente, pela Declaração de Berlim, o movimento de acesso aberto estabeleceu duas importantes estratégias, definidas como linha verde e dourada. A primeira refere-se “[...] ao depósito de trabalhos científicos que tenham sido publicados em revistas de acesso restrito em repositórios temáticos ou institucionais [...]” (OLIVEIRA; GOMES, 2013, p. 2). Segundo os mesmos autores, essa estratégia também pode ser chamada de autoarquivamento, pois é realizada pelos próprios autores dos artigos. Já a segunda diz respeito “[...] aos trabalhos científicos publicados em revistas que aderiram ao Movimento de Acesso Aberto ou Open Access (OA)” (OLIVEIRA; GOMES, 2013, p. 2).

No caso específico da disponibilização de periódicos científicos em espaço aberto, segundo Ortellado (2008 apud OLIVEIRA; GOMES, 2013), destaca-se o grande obstáculo para a legitimação desse movimento, pois, o processo editorial realizado muitas vezes por instituições privadas é comprometido, devido à democratização dos documentos antes disponibilizados de forma comercial.

A respeito da importância do acesso aberto a informações científicas e em meio às diversas discussões existentes, Kuramoto (2009, p. 7) disserta que

O acesso à informação científica tornou-se, em consequência das barreiras existentes, um dos grandes desafios no mundo de hoje. Uma dessas barreiras, o custo crescente da assinatura dos principais periódicos científicos, provocou a chamada crise dos periódicos científicos. Para superar essa crise, pesquisadores de diversas partes do globo terrestre se reuniram e deram início a um grande movimento global em direção ao acesso aberto à informação científica.



Os defensores do movimento de acesso aberto partem do princípio de que se deve priorizar a oferta gratuita de informações, concretizando a democratização do conhecimento, sem limitações. Defende-se a eliminação de quaisquer barreiras financeiras, legais ou técnicas, além daquelas próprias do acesso à *internet*. A única restrição à reprodução e distribuição e a única função do *copyright* nesse contexto devem ser o controle dos autores sobre a integridade de sua obra e o direito de serem adequadamente reconhecidos e citados (BUDAPEST OPEN ACCESS INITIATIVE, 2001 apud LEITE, 2009).

Se por um lado, o movimento do acesso aberto viabiliza o processo de disseminação da informação, tornando o conhecimento mais acessível, em contrapartida, surge a necessidade de se desenvolver mecanismos que garantam o respeito aos responsáveis pela criação e divulgação da informação, atendendo a princípios legalmente instituídos por legislações específicas (FREITAS; MAIA; LEITE, 2011).

É perceptível que esses processos são complexos, o que requer o estabelecimento de políticas voltadas para gerenciar e controlar com integridade e legitimidade as informações disponibilizadas. A pesquisa de Ferrari e Pires (2014) sinaliza essa necessidade, enfatizando haver prioridade na elaboração de diretrizes que promovam a normatização, tendo em vista, o uso adequado do compartilhamento de informações, sem provocar infrações à legislação em vigor.

As complexidades advindas com o uso e expansão acentuada da *internet* condicionam a elaboração de parâmetros que garantam a utilização legitimada do ciberespaço, tornando-o um ambiente lícito e influenciador na construção de conhecimento (SANTA ANNA, 2013).

Por conseguinte, instituir regras constitui um trabalho cauteloso que requer pesquisas prévias, a fim de garantir maior consistência e integridade nas normas, instrumentos e estratégias de controle. Assim, antes de qualquer ação, é preciso entender o trabalho de organização da informação e as complexidades do ciberespaço, com vistas a construir uma cultura colaborativa, levando em consideração inúmeras questões, sobretudo os interesses que envolvem as decisões do poder público (FERRARI; PIRES, 2014).



MATERIAIS E MÉTODOS

A partir das questões refletidas por teóricos renomados da área de Biblioteconomia e Ciência da Informação, constata-se que, a utilização do ambiente virtual, especialmente com o desenvolvimento da internet, desencadeou muitos desafios que precisam ser superados a fim de garantir um uso efetivo da informação socializada pelos novos artefatos tecnológicos.

O movimento do acesso aberto, particularmente, evidencia problemas ainda mais elevados, pois além das questões complexas envolvidas no processo de organização e representação da informação eletrônica, remete ao estabelecimento de padrões normativos a fim de que a informação seja utilizada de forma lícita, sem provocar infrações contra a propriedade intelectual e editorial.

Partindo dessa contextualização, este estudo, realizado de forma preliminar, investiga em campo como um grupo de usuários de informação eletrônica, utiliza as informações disponibilizadas em ambiente digital, tendo em vista as complexidades envolvidas no movimento do acesso aberto.

Foram escolhidos como amostragem de pesquisa, três pesquisadores, sendo um aluno de graduação (entrevistado A), um aluno de mestrado (entrevistado B) e um aluno de doutorado (entrevistado C). Também compuseram a amostragem, dois profissionais envolvidos no mercado editorial da informação, sendo um bibliotecário atuante na organização de repositório de dados científicos (entrevistado D), e, por fim, um editor de revista eletrônica (entrevistado E). Na escolha desses sujeitos, levou-se em conta os principais indivíduos envolvidos no processo de pesquisa no âmbito universitário.

Os métodos de pesquisa empregados foram a pesquisa bibliográfica, a documental e estudo de campo. A pesquisa do tipo bibliográfica, segundo Gil (2010, p. 41), tem como objetivo [...] o aprimoramento de ideias ou descoberta de instituições. Seu planejamento é, portanto, bastante flexível, de modo que possibilite a consideração dos variados aspectos ao fato estudado [...]. Os materiais utilizados nesse método foram livros e artigos científicos publicados em revistas da área de Ciência da Informação.

Já a pesquisa documental, segundo o mesmo autor, emprega os materiais que ainda não sofreram tratamento de análise, ou que possam ainda ser reestruturados. As fontes utilizadas nesse método foram fontes de aspecto jurídico, como a Constituição Federal e a Lei de Direitos Autorais (Lei nº 9.610).

O estudo de campo foi realizado em uma instituição de ensino superior (universidade), tendo como sujeitos de pesquisa, três pesquisadores e dois profissionais atuantes na disponibilização de informação eletrônica, todos atuantes nessa instituição. O estudo de campo representa uma investigação, cujo objetivo é estudar a realidade dos indivíduos, grupos, comunidades e instituições (GIL, 2010). A técnica utilizada na coleta de dados foi a entrevista aberta, cujo objetivo é obter informações de interesse a uma investigação, onde o pesquisador formula perguntas orientadas, com um objetivo definido, frente a frente com o respondente e dentro de uma interação social (GIL, 1999).



Quanto aos procedimentos para apresentação e análise dos dados, foi utilizada a descrição fiel das falas dos respondentes, analisando-as de acordo com as recomendações propostas pela literatura e por legislação vigente.

APRESENTAÇÃO E ANÁLISE DE DADOS

Primeiramente, os três sujeitos pesquisadores foram indagados a respeito de quais as fontes de informação mais utilizadas na condução de suas pesquisas, enfatizando, de modo especial, o uso de fontes eletrônicas, disponibilizadas na internet.

Os entrevistados foram unânimes ao destacarem que utilizam fontes impressas, recorrendo à biblioteca universitária da instituição, a fim de conseguir livros e artigos impressos, sobretudo quando precisam realizar estudos mais aprofundados. No entanto, a maioria dos materiais empregados são artigos publicados em revistas eletrônicas da área de interesse.

“[...] Utilizo a biblioteca poucas vezes, pois com a internet, consigo baixar artigos recentes, realizando a leitura, sem exigir grande esforço tanto na localização dos materiais quanto no momento de citar a fonte. Se a fonte me atende, compartilho com outros colegas de interesse [...]” (Entrevistado A).

“[...] A *internet* facilita meu trabalho, pois muita coisa já vem pronta, como as referências, bastando apenas copiar. [...] Mas, mesmo assim, ainda vou à biblioteca a procura de livros” (Entrevistado B).

“[...] Utilizo muito livro da biblioteca ou comprando-os [...] Os artigos consigo através de bases de dados confiáveis existentes no Portal de Periódicos da Capes, mesmo tendo algumas dificuldades para utilizá-lo, devido às muitas possibilidades oferecidas por esse espaço digital” [...] (Entrevistado C).

Nesse mesmo enfoque, os dois profissionais que trabalham na indústria editorial eletrônica foram investigados a respeito dos impactos que o uso das fontes eletrônicas vem desencadeando, sobretudo ao trabalho dos profissionais atuantes nesse mercado. De modo geral, esses sujeitos afirmaram que, a informação disponibilizada na *internet* facilita o processo de pesquisa dos usuários, no entanto, muitos problemas são originados devido ao uso indevido e/ou inadequado do espaço digital.

“[...] A criação da *internet* e seu uso na disponibilização de informações foi um dos maiores avanços [...]. Só que, para que a informação seja recuperada, precisa ser feita uma efetiva representação dos objetos digitais [...] Também é preciso capacitar o usuário para que aprenda a usar de modo adequado, de modo que explore todas as ferramentas disponíveis, bem como não infrinja leis de direitos autorais [...]” (Entrevistado D).

“A *internet* facilitou a vida das pessoas [...] Vejo que os profissionais que atuam nessa área devem disponibilizar a informação, preocupando-se em estabelecer estratégias que eduque os clientes ao uso legal da informação disponibilizada [...] Os editores, em especial, devem reformular suas práticas a fim de conquistar esse novo espaço de trabalho que está aflorando [...]” (Entrevistado E).



Quanto a essa questão do novo ambiente existente para disponibilização e uso da informação científica, constata-se, com base na realidade estudada e com opiniões de autores da área, que esse espaço provoca inúmeros benefícios aos usuários da informação. Esses benefícios referem-se tanto aos processos de troca e busca de informações, quanto ao comportamento do usuário ao adquirir maior autonomia (OLIVEIRA; SANTOS; RODEGHERI, 2013).

Importante enfatizar que o advento da *internet* emerge como uma forma de efetivar a garantia dos direitos do cidadão no que se refere ao acesso à informação, tornando-a mais democrática. Essa garantia é fruto das possibilidades de manifestação e divulgação de ideias, através da livre circulação de conteúdo na *web*. No entanto, é necessário estabelecer medidas de controle, a fim de evitar conflitos no ciberespaço (OLIVEIRA; SANTOS; RODEGHERI, 2013).

Após essa discussão a respeito do uso das informações eletrônicas, os pesquisadores foram questionados a respeito do conhecimento que possuem em relação ao uso das tecnologias digitais e o acesso a documentos livres de forma lícita. Em linhas gerais, constatou-se que os três pesquisadores não conhecem muitas potencialidades oferecidas, assim como desconhecem as medidas impostas para garantia de direitos autorais e editoriais.

“[...] Localizo os artigos por meio da busca no Google, mesmo [...] Todo artigo que aparece, se for pertinente para o assunto que estou tratando, copio as informações e cito a fonte, sem preocupar-me com local que foi publicado ou outras considerações” [...] (Entrevistado A).

“[...] Não gosto de procurar informações por meio dos catálogos das bases de dados, pois os acho confusos [...]. Vejo a *internet* como um espaço público para acessar e compartilhar informação. Sendo assim, não prendo-me a questões se a fonte tem algum direito reservado ou licenças específicas. O importante é citar a fonte” [...] (Entrevistado B).

“[...] As ferramentas e estratégias para buscar a informação eletrônica são bacanas. Ajudam muito na localização, no entanto, muitas vezes, a informação existe na base de dados e não consegue ser localizada [...]”. Eu penso que os direitos autorais são preservados, mas os direitos editoriais não tem como garanti-los, uma vez que a *internet* é um espaço democrático de compartilhamento [...]” (Entrevistado C).

A respeito do que fazem os profissionais no intuito de tornar as informações mais acessíveis para o usuário, viabilizando com mais intensidade a disseminação, os profissionais consideram como necessária a realização de um trabalho interdisciplinar, devendo considerar, *a priori*, as necessidades, dificuldades e costumes dos usuários, a fim de estabelecer melhores interfaces e formas de representação, assim como a construção de instrumentos normativos a fim de evitar o uso inadequado das informações existentes em acesso aberto.



“[...] A construção do repositório dessa biblioteca reflete um trabalho árduo que não dá para ser feito isoladamente, mas requer a presença de vários profissionais [...] é preciso que seja valorizada a formação de uma comissão para tomar decisões a respeito das práticas de tratamento da informação, e principalmente fazer estudos de usuários a fim de conhecer o perfil da comunidade usuária (Entrevistado D).

“[...] Disponibilizamos serviço de sugestões no sítio da revista e com base nessas sugestões melhoramos a interface da revista [...] Temos o cuidado de disponibilizar os metadados da publicação, bem como *Copyright, creative commons*, dentre outras licenças e direitos relacionados à proteção intelectual [...] (Entrevistado E).

Com base nessas discussões, entende-se que, os profissionais da informação atuantes na construção de repositórios ou base de dados científicos devem seguir recomendações específicas da instituição à qual pertençam esses instrumentos. Essas recomendações “[...] constituem um conjunto de instruções que sistematizam processos que devem ser considerados no momento da elaboração e execução de um projeto de repositório institucional em universidades e institutos de pesquisa” (LEITE, 2009, p. 13).

A utilização de catálogos dinâmicos, permeados por diversas estratégias de busca e recuperação da informação, denominados de catálogo em linha de acesso público (OPAC), representou um grande avanço na localização de informações, diante da imensa quantidade de informação disponibilizada em ambientes eletrônicos. No entendimento de Oliveira (2008), a grande vantagem é a capacidade simultânea de recuperação da informação, além da variada possibilidade de pontos de acesso, por meio da introdução de hipertextos, viabilizando buscas não lineares e dinâmicas.

Quanto à preservação de direitos autorais e editoriais, analisando a legislação específica de direitos autorais, Lei nº 9610, instituída no Brasil no ano de 1998, é preciso considerar que, mesmo estando em ambiente eletrônico, esses direitos devem ser respeitados. O art. 7º dessa lei define que as obras intelectuais protegidas são as criações do espírito, expressas por qualquer meio ou fixadas em qualquer suporte, tangível ou intangível, conhecido ou que se invente no futuro, tais como: os textos de obras literárias, artísticas ou científicas; as conferências, alocações, sermões e outras obras da mesma natureza; as obras dramáticas e dramático-musicais; as obras coreográficas e pantomímica, dentre outros.

Vê-se que os direitos autorais também são garantidos no ciberespaço, cabendo aos profissionais, elaborar instrumentos normativos que assegurem a efetivação desses direitos. Além disso, esses profissionais devem estabelecer ações voltadas para a educação e conscientização dos usuários quanto ao respeito a esses direitos. É importante refletir que, o controle não deve impedir o acesso, uma vez que, na *internet* amplia-se o espaço para a efetivação dos direitos constitucionais da liberdade de expressão e informação, pois se permite a livre emissão do pensamento, de forma célere, transpondo barreiras temporais e territoriais (OLIVEIRA; SANTOS; RODEGHERI, 2013).



De acordo com a Lei Maior da nação, o acesso à informação representa um dos direitos e garantias fundamentais do cidadão, desde que esse acesso seja realizado de forma lícita, sem infringir normas previamente estabelecidas, garantindo direitos de propriedade intelectual sobre as origens das informações (BRASIL, 1988).

A partir do exposto, entende-se que os profissionais que atuam na indústria editorial eletrônica, sobretudo os bibliotecários devem ser catalisadores e disseminadores da informação, bem simbólico e elemento chave na conquista da cidadania, mas também educadores, pois a biblioteca e demais instituições de informação constituem um *locus* privilegiado do debate, esferas de discussão política e ambientes responsáveis pela democratização da informação (MORIGI; VANS; GALDINO, 2002).

Visando conhecer as maiores dificuldades dos usuários da informação eletrônica, suas percepções e indicações de melhoria, os pesquisadores foram indagados sobre o que consideravam como maiores obstáculos, facilidades e sugestões a serem implementadas, tendo em vista, garantir facilidade do uso das ferramentas, efetividade no processo de busca e legitimidade na utilização das fontes.

“O maior obstáculo que me afeta é encontrar muitos *links* desativados, documentos que não abrem e também documentos que não estão disponíveis para download [...]. Acho bom a capacidade de navegarmos de um local para outro, ao mesmo tempo [...] Seria bom tornar os sites mais agradáveis, com figuras e letras mas destacadas [...]” (Entrevistado A).

“[...] Tenho dificuldade em acessar certas bases de dados que pedem cadastro, número de senha, e às vezes esse acesso somente é permitido enquanto estamos nas dependências da universidade [...]. Sugiro que essas dificuldades de acesso fossem mais facilitadas [...]” (Entrevistado B).

“[...] Acho fantástico algumas estratégias de busca criadas, bem como os indicadores de citações, [...] Tenho algumas dificuldades em utilizar todas essas possibilidades [...] O acesso muitas vezes é restrito a certos locais [...] Como melhoria, sugiro a realização de treinamentos, disponibilização de tutorias no *site*, divulgação dos treinamentos [...]” (Entrevistado C).

Sobre essas questões acima apresentadas, os profissionais foram indagados a respeito do que tentam fazer para reduzir as problemáticas geradas com o uso intensificado da informação eletrônica, tornando essas práticas cada dia mais utilizáveis, democráticas e seguras.

“[...] Os problemas são grandes. As dúvidas também. É preciso participação da universidade e do poder público no sentido de adotar ações voltadas para a resolução dessas problemáticas [...] Entendo que a educação do usuário é a principal estratégia para atingir o sucesso, pois percebemos que muitos produtos e serviços oferecidos não são utilizados [...] Realizamos poucos estudos de usuários e poucos treinamentos com essa especificidade [...]” (Entrevistado D).

“[...] Imagino que a questão da comodidade oferecida nas interfaces dos periódicos certamente é um fator determinante na satisfação dos leitores [...] Aqui, nós temos apoio de profissionais para tornar o *site* mais atrativo [...] Mas, é preciso firmar outras parcerias, especialmente com bibliotecários [...] A questão legal que envolve o uso da informação seria melhor controlada se fossem melhores divulgadas na sociedade, as questões legais [...]” (Entrevistado E).



A respeito do que consideram os pesquisadores entrevistados, discute-se, com grande frequência na literatura, como a arquitetura informacional interfere na visibilidade. De acordo com Gruszynski e Golin (2007), é preciso utilizar novos formatos de disponibilização de metadados e práticas de representação mais consistentes. De modo geral, é prioritário sintonizar as informações disponíveis, tornar a mensagem única e assegurar a credibilidade do que está sendo comunicado.

No que se refere às problemáticas relacionadas a direitos autorais e demais licenças, não resta dúvida de que a educação pode representar uma estratégia viável, com vistas a consolidar a formação consciente no uso da informação. Sobre a capacitação do usuário, através de práticas educativas como os treinamentos, o papel de bibliotecário-educador será tanto na instrução formal para o uso das tecnologias de informação quanto no planejamento de softwares interativos, ou outras formas de comunicação para o acesso à informação (CUENCA, 1999).

Dentre os vários benefícios que a educação/capacitação do usuário pode representar, destaque especial pode ser atribuído ao desenvolvimento de buscas na sua área, na seleção e localização da informação, no conhecimento das várias formas de acesso à informação e ao documento, e, por fim, a tentativa de manter o usuário constantemente atualizado nessas tecnologias e formas de acesso (CUENCA et al., 1999).



RESULTADOS PARCIAIS/FINAIS

Seguindo os objetivos propostos na condução desta pesquisa, e após apresentação e discussão dos dados coletados, é possível obter alguns resultados parciais, com base na realidade analisada e tendo as duas categorias de sujeitos investigados, pesquisadores e profissionais da informação. Abaixo estão elencados os resultados alcançados:

- a produção científica na atualidade é viabilizada por meio do uso de fontes disponibilizadas em ambiente virtual; os pesquisadores analisados utilizam essas fontes com mais intensidade do que materiais impressos encontrados nos acervos de bibliotecas;
- os pesquisadores entrevistados enfrentam muitos obstáculos ao fazerem uso das informações eletrônicas, sendo os principais: problemas de estruturação dos instrumentos de busca, fontes desativadas, conteúdo parcial das informações, não reconhecimento de licenças e direitos autorais e a utilização de informações não confiáveis e inseguras;
- como vantagens principais no acesso à informação eletrônica os pesquisadores reconhecem: a criação de instrumentos facilitadores na busca e recuperação da informação, a utilização dos OPACs, a rapidez de localização e compartilhamento da informação, interatividade e o acesso a publicações recentes. Como melhorias, sugerem a reestruturação da arquitetura informacional e a realização de treinamentos a fim de capacitá-los quanto ao acesso e uso da informação no ambiente digital;
- os profissionais destacam o valor que as informações eletrônicas podem proporcionar à produção científica, porém enfatizam problemas relacionados aos direitos autorais.
- os profissionais sinalizam como possibilidades de solução, o trabalho colaborativo e interdisciplinar entre diferentes áreas, assim como o engajamento e valorização por parte da universidade e do poder público em formular ações voltadas para a educação da população, tornando-a capacitada a usufruir das inúmeras possibilidades oriundas com o uso do ambiente digital.
- A educação do usuário viabilizada pela prática do treinamento torna-se uma estratégia de duplo benefício pois, além de capacitar o usuário quanto ao uso dos novos artefatos tecnológicos, também torna-o ciente quanto ao uso lícito de licenças e direitos relativos à propriedade intelectual e editorial.



CONSIDERAÇÕES PARCIAIS/FINAIS

Com base nas discussões propostas e apresentadas neste estudo, constatou-se que o acesso às informações eletrônicas, potencializado com a origem da internet, produziu inúmeros impactos em toda a ambiência que perpassa a produção científica e a indústria editorial. Este fato desencadeia a necessidade de reestruturar todo o processo de comunicação científica, envolvendo novas práticas a serem utilizadas por pesquisadores, editores e profissionais da informação.

O paradigma da disseminação da informação, fortemente instituído na sociedade contemporânea, aliado à consolidação do ciberespaço, remete-se ao movimento do acesso aberto. Embora desperte inúmeras vantagens, esse movimento requer o estabelecimento de padrões normativos a fim de evitar práticas ilícitas por parte dos usuários no uso da informação, além de condicionar um novo espaço de trabalho para os profissionais envolvidos com a transferência da informação.

O estudo constatou, com base em entrevista aplicada a pesquisadores e profissionais da informação de uma universidade, que o novo espaço de comunicação científica proporciona inúmeras vantagens, sobretudo àquelas relacionadas à localização e compartilhamento da informação, sem limitações de tempo e de espaço. No entanto, também destacam-se alguns impactos, especialmente às complexidades dos objetos digitais e sua disponibilização em repositórios e bases de dados, no intuito de garantir segurança e legalidade quanto ao uso desses objetos sem provocar infrações à legislação vigente sobre direitos autorais e editoriais.

Com base neste estudo, conclui-se que é preciso estabelecer políticas de acesso a esses ambientes, bem como a divulgação de como utilizar os documentos sem impactar com problemas de direitos autorais. Também, constatou-se, tanto por parte dos pesquisadores quanto profissionais, a necessidade de capacitar o usuário da informação eletrônica, tendo em vista sua adequação aos limites impostos pelo ambiente virtual.

As discussões relacionadas ao uso da informação em acesso aberto trazem à baila inúmeros aspectos relevantes, despertando a necessidade da elaboração de novos estudos que indiquem soluções para tornar essas práticas mais seguras e lícitas, condicionando a democratização da informação, com vistas à garantia de direitos e deveres de todos os cidadãos inseridos na sociedade da informação.

Assim, recomenda-se a realização de pesquisas junto a profissionais especializados em tecnologia da informação, a fim de elaborar novos instrumentos, com vistas a garantir melhor representação e recuperação da informação. Também constatou-se a necessidade de estudos interdisciplinares com profissionais do Direito, tendo em vista elaborar medidas normativas que assegure o respeito à propriedade intelectual e editorial, garantindo a concretização dos direitos e deveres de todos os envolvidos no mercado da informação.



REFERÊNCIAS BIBLIOGRÁFICAS

BRASIL. **Constituição da República Federativa do Brasil (1988)**. Brasília: Senado Federal, 1988.

_____. **Lei nº 9.610**, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l9610.htm>. Acesso em: 5 ago. 2014.

CASTELLS, Manuel. **A sociedade em rede**. 10. ed. rev. ampl. São Paulo: Paz e Terra, 2007.

CUENCA, Angela Maria Belloni. O usuário final da busca informatizada: avaliação da capacitação no acesso a bases de dados em biblioteca acadêmica. **Ci. Inf.**, Brasília, v. 28, n. 3, p. 293-301, set./dez. 1999. Disponível em: <<http://www.scielo.br/pdf/ci/v28n3/v28n3a7.pdf>>. Acesso em: 6 ago. 2014.

_____; et al. Capacitação no uso das bases Medline e Lilacs: avaliação de conteúdo, estrutura e metodologia. **Ci. Inf.**, Brasília, v. 28, n. 3, p. 340-346, set./dez. 1999. Disponível em: <<http://www.scielo.br/pdf/ci/v28n3/v28n3a12.pdf>>. Acesso em: 5 ago. 2014.

DELEUZE, Gilles; GUATTARI, Félix. **Mil platôs: capitalismo e esquizofrenia**. Rio de Janeiro: Ed. 34, 1995.

FERRARI, Rodrigo Duarte; PIRES, Giovani De Lorenzi. Auto-arquivamento e acesso aberto: deveres e direitos digitais na sociedade de rede. **Rev. digit. bibliotecon. cienc. inf.** Campinas, SP, v. 12, n. 1, p. 22-38, jan./abr. 2014. Disponível em: <http://www.sbu.unicamp.br/seer/ojs/index.php/rbci/article/view/3874/pdf_48>. Acesso em: 5 ago. 2014.

FREITAS, Marília Augusta; MAIA, Luanna Cezar; LEITE, Fernando César Lima. Acesso aberto como estratégia de disseminação e preservação da produção científica discente: a Biblioteca Digital de Monografias da Universidade de Brasília. **Bibl. Univ.**, Belo Horizonte, v. 1, n. 1, p. 71-80, jan./jun. 2011. Disponível em: <<file:///C:/Users/Guest/Downloads/24-170-1-PB.pdf>>. Acesso em: 5 ago. 2014.

GIL, Antonio Carlos. Entrevista. In: _____ **Métodos e Técnicas de Pesquisa Social**. 5. ed. São Paulo: Atlas, 1999. Cap. 11, p.117-127.

_____. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.



GRUSZYNSKI, Ana Cláudia; GOLIN, Cida. Periódicos científicos eletrônicos e a visibilidade da ciência na web: estudo de caso na UFRGS. **DataGramZero**. Rio de Janeiro, v. 8, n 3, jun. 2007. Disponível em: <http://www.dgz.org.br/jun07/Art_02.htm>. Acesso em: 6 ago. 2014.

KURAMOTO, Hélio. Prefácio. In: LEITE, Fernando César Lima. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto. Brasília: IBICT, 2009, p. 7-10.

LEITE, Fernando César Lima. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto. Brasília: IBICT, 2009.

LÉVY, Pierre. **Cibercultura**. São Paulo: Editora 34, 1999.

MOREGI, Valdir José; VANZ, Samile Andréa de Souza; GALDINO, Karina. O bibliotecário e suas práticas na construção da cidadania. **Revista ACB**, Vol. 7, No 2, 2002. Disponível em: <http://revista.acbsc.org.br/index.php/racb/article/viewArticle/390/479>. Acesso em: 5 de ago. 2014.

MUELLER, Suzana Pinheiro Machado. A ciência, o sistema de comunicação científica e a literatura científica. In: CAMPELLO, Bernadete Santos; CENDÓN, Beatriz Valadares; KREMER, Jeannette Marguerite. **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: Ed. UFMG, 2000, p. 21-34.

_____. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ci. Inf.**, Brasília, v. 35, n. 2, p. 27-38, maio/ago. 2006. Disponível em: <<http://www.scielo.br/pdf/ci/v35n2/a04v35n2.pdf>>. Acesso em: 5 ago. 2014.

OLIVEIRA, Carla Cristina Vieira de. A interação de usuários da UFMG com o catálogo on-line do sistema do Pergamum . **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 4, n. 2, p. 73-88, jul./dez. 2008. Disponível em: <http://www.febab.org.br/rbbd/ojs-2.1.1/index.php/rbbd/article/viewFile/104/144>. Acesso em: 5 ago. 2014.

OLIVEIRA, Flásleandro Vieira Oliveira; GOMES, Sandra Lúcia Rebel Gomes. A licença Creative Commons e o movimento de acesso aberto à informação científica: um olhar para a Scientific Electronic Library on Line (SciELO). In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 25., 2013, Florianópolis. **Anais eletrônicos...** Florianópolis: FEBAB, 2013. p. 1-16. Disponível em: <file:///C:/Users/aluno-ccje/Desktop/1602-1615-1-PB.pdf>. Acesso em: 2 out. 2014.



OLIVEIRA, Rafael Santos de Oliveira; SANTOS, Noemi de Freitas; RODEGHERI, Letícia Bodanese. Judicialização de conflitos no ciberespaço: desafios à liberdade de expressão na blogosfera. **Revista de Direitos Fundamentais e Democracia**, Curitiba, v. 13, n. 13, p. 160-178, jan./jun. 2013.

PACKER, Abel Laerte. A construção coletiva da Biblioteca Virtual em Saúde. **Interface**, v.9, n.17, p.249-72, mar./ago. 2005. Disponível em: <<http://www.scielo.br/pdf/icse/v9n17/v9n17a04.pdf>>. Acesso em: 5 ago. 2014.

SANTA ANNA, Jorge. A (r)evolução digital e os dilemas para a catalogação: os cibertecários em atuação. In: ENCONTRO INTERNACIONAL DE CATALOGADORES, 9; ENCONTRO NACIONAL DE CATALOGADORES, 2, **Anais Eletrônicos**. Rio de Janeiro: Biblioteca Nacional, 2013. Disponível em: <[file:///C:/Users/aluno-ccje/Downloads/21-185-1-PB%20\(2\).pdf](file:///C:/Users/aluno-ccje/Downloads/21-185-1-PB%20(2).pdf)>. Acesso em: 15 abr. 2014.

WEITZEL. S. R. O papel dos repositórios institucionais e temáticos na estrutura da produção científica. **Em Questão**, v. 12, n. 1, p. 51-71, 2006. Disponível em: <http://eprints.rclis.org/12101/1/weitzel_repositorios.pdf>. Acesso em: 5 ago. 2014.



APÊNDICE A – ROTEIRO DE PERGUNTAS PARA ENTREVISTA

PERGUNTAS DIRECIONADAS AOS PESQUISADORES

- 1 – Ao realizarem suas pesquisas, em quais locais vocês mais recorrem? Quais fontes de informação vocês mais utilizam? Utilizam as fontes de pesquisa disponibilizadas na internet?
- 2 – Vocês possuem conhecimento necessário para utilizar as potencialidades oferecidas pelas novas tecnologias? Sabem selecionar as informações confiáveis e disponibilizadas para uso? Conhecem as leis que regem a utilização/acesso das informações eletrônicas?
- 3 - O que vocês consideravam como maiores obstáculos, facilidades e sugestões a serem implementadas, tendo em vista, garantir facilidade do uso das ferramentas, efetividade no processo de busca e legitimidade na utilização das fontes eletrônicas?

PERGUNTAS DIRECIONADAS AO BIBLIOTECÁRIO E EDITOR

- 1 – Quais os impactos que o uso das fontes eletrônicas vem desencadeando, sobretudo ao trabalho dos profissionais atuantes no mercado da indústria editorial?
- 2 – No entendimento de vocês, o que pode ser feito para que as informações disponibilizadas no ambiente digital sejam mais acessíveis para o usuário, tendo em vista a disseminação da informação?
- 3 – O que vocês tentam fazer para reduzir as problemáticas geradas com o uso intensificado da informação eletrônica, tornando essas práticas cada dia mais utilizáveis, democráticas e seguras?



LETRAS ESPANHOL NO REPOSITÓRIO INSTITUCIONAL DE CONTEÚDO DIGITAL UFSC

Morena Pereira Porto¹, Paula Balbis Garcia², Vera Regina de Aquino Vieira³

¹ Graduanda do curso de Biblioteconomia na Universidade Federal de Santa Catarina (UFSC). Estagiária e responsável pela coleção de Objetos de Aprendizagem do curso de Letras Espanhol EaD no Repositório UFSC.

² Graduação em Licenciatura em Letras Língua Espanhola e Língua Portuguesa e respectivas literaturas (2000) e Mestrado em Ciência da Informação (2014) pela Universidade Federal de Santa Catarina. Designer Instrucional do curso de Letras Espanhol na modalidade a distância e Tutora da disciplina metodologia do ensino de línguas estrangeiras/adicionais pelo mesmo curso.

³ Graduação em Letras Português e Espanhol pela Universidade Federal de Santa Catarina (1985). Especialização em Curso para formação de professores de Español LE pela Universidade Complutense de Madrid (1986) e Doutorado em Filología Española I pela mesma Universidade. Professora da Universidade Federal de Santa Catarina, no Departamento de Língua e Literatura Estrangeiras, atuando principalmente nos seguintes temas: Espanhol LE, linguística, Linguística Aplicada, Aquisição, Leitura, Texto e Pragmática.

RESUMO

Com a Internet e as facilidades de acesso a recursos colaborativos da web social, os usuários dispõem de autonomia e agilidade no acesso às informações. A EaD acontece por meio das tecnologias digitais e esses recursos auxiliam a inovar a metodologia educacional tradicional, favorecendo a criação de materiais multimodais para realização de estudos. Devido à grande quantidade de conteúdos digitais produzidos para a primeira edição do curso de graduação em Licenciatura em Letras Espanhol EaD, da Universidade Federal de Santa Catarina (UFSC), os tutores e professores atuantes no curso sentiram a necessidade de localizar e recuperar de maneira mais rápida os Objetos de Aprendizagem (OAs) das disciplinas dos semestres anteriores. Desta forma, a Coordenação, juntamente com a equipe de produção de material didático optou pela gestão e disseminação dos materiais em um único local, que fosse institucional e aberto a toda a comunidade universitária e demais interessados. Neste caso, o Repositório Institucional de Conteúdos Digitais da UFSC foi pensado como sendo a ferramenta mais propícia para a padronização, organização e o compartilhamento dos objetos, já que era institucional, gratuito e de livre acesso, além de ter como proposta tornar disponível e acessível a produção científica da UFSC. Portanto, o objetivo deste trabalho é o relato da experiência e da metodologia utilizada para o mapeamento, seleção, organização e disseminação dos OAs produzidos no curso de Língua Espanhola EaD no Repositório UFSC. Para a realização deste estudo de caráter tanto descritivo quanto exploratório, foram definidos diferentes métodos de pesquisa ao longo do processo, como o desenvolvimento de metodologias de tratamento da informação e levantamento de dados. Apesar de ainda não finalizado, as investigadoras apresentam



consideráveis resultados a serem ampliados e analisados. A primeira etapa do trabalho, ainda em andamento, totaliza 1007 OAs mapeados das duas primeiras edições, sendo 207 já indexadas no Repositório UFSC até o presente momento. Para avaliar o impacto das ações tomadas na etapa inicial do trabalho, foi enviado um questionário on-line direcionado a 49 profissionais do curso (tutores e professores) sobre a utilização do Repositório UFSC, levando em conta o processo e a capacitação dos usuários. Por meio do questionário, foram obtidas 15 respostas com informações pertinentes a fim de auxiliar no acompanhamento do trabalho feito, permitindo a reflexão sobre as práticas e decisões tomadas bem como a retroalimentação das etapas anteriores.

Palavras-chave: Repositório Institucional; Objetos de Aprendizagem; Educação a Distância; Acesso Aberto

LETRAS ESPAÑOL EN EL REPOSITÓRIO INSTITUCIONAL DE CONTEÚDO DIGITAL UFSC

RESUMEN

Con la Internet y las facilidades de acceso a recursos colaborativos de la web social, los usuarios disponen de autonomía y más velocidad en el acceso a las informaciones. En la educación, esos recursos dan soporte para la innovación de la metodología educacional tradicional, como, por ejemplo, en la modalidad de enseñanza a distancia. La EaD ocurre por medio de tecnologías digitales, y esos recursos corroboran con la innovación de la metodología educacional tradicional, favoreciendo la creación de materiales multimodales para la realización de estudios. Debido a la gran cantidad de contenido digital producido para la primera edición del curso de graduación en Licenciatura en Letras Español de la *Universidade Federal de Santa Catarina* (UFSC), los tutores y profesores del curso sintieron la necesidad de localizar y recuperar de manera más simple y rápida los Objetos de Aprendizaje (OAs) de las asignaturas de los semestres anteriores. De esta forma, la Coordinación y el equipo de producción de material didáctico decidieron por la gestión y disseminación de los materiales en un único lugar, que fuera institucional y abierto a toda la comunidad universitaria y demás interesados. En este caso, el *Repositório Institucional de Conteúdos Digitais da UFSC* fue la herramienta más propicia para la patronización, organización y compartimiento de los OAs, ya que es gratuito y de acceso libre, además de tener como propuesta tornar disponible y accesible la producción científica de la UFSC. El objetivo de este trabajo es el relato de la experiencia y de la metodología utilizada para el mapeo, selección, organización y disseminación de los OAs producidos en el curso de Lengua Española a distancia en el Repositorio UFSC. Para la realización de este estudio de carácter tanto descriptivo como exploratorio, fueron definidos distintos métodos de investigación a lo largo del proceso, como el desarrollo de metodologías del tratamiento de la información y el levantamiento de datos. Pese a no estar finalizado, las investigadoras presentan,



en este trabajo, considerables resultados que todavía serán ampliados y analizados. La primera instancia del trabajo, todavía inacabado, totaliza 1007 OAs mapeados de las dos primeras ediciones del curso, siendo que, de estas, 207 ya fueron indexadas en el Repositorio UFSC hasta el presente momento. Para evaluar el impacto de las acciones tomadas en la etapa inicial del trabajo, fue enviada una encuesta en línea para los 49 profesionales del curso (tutores y profesores) sobre la utilización del Repositorio, teniendo en cuenta el proceso y capacitación de los usuarios. Con la encuesta, fueron obtenidas 15 respuestas con informaciones pertinentes a ser discutidas con el fin de auxiliar en la orientación del trabajo hecho, permitiendo la reflexión sobre las prácticas y decisiones tomadas, así como el *feedback* de las etapas anteriores.

Palabra clave: Repositorio Institucional; Objetos de Aprendizaje; Educación a distancia; Acceso Abierto

INTRODUÇÃO

O constante progresso e desenvolvimento tecnológico vivido na sociedade da informação influenciam e até reestruturam as práticas e relações na educação, criando conseqüentemente novas formas de pensar e agir sobre esta. Com a Internet e as facilidades de acesso a recursos colaborativos da web social, os usuários dispõem de mais autonomia e agilidade no acesso às informações. Na educação, esses recursos auxiliam a inovar a metodologia educacional tradicional e, por conseqüência, o aluno ao ter domínio sobre as ferramentas da web, assume um novo papel, sendo não mais um receptor, mas um ator proativo na comunidade escolar, tornando-se responsável por sua aprendizagem. Gazzoni (2006) afirma que o computador passou a ser um meio de ensino e tem acelerado o desenvolvimento do aprendizado eletrônico (e-learning) possibilitando a utilização de conteúdos didáticos disponibilizados na rede em diferentes formatos e modalidades de ensino, como por exemplo, a Educação a Distância (EaD).

A EaD é uma modalidade de educação mediada pela tecnologia na qual alunos e professores estão separados espacial e/ou temporalmente, ou seja, não estão fisicamente presentes em um mesmo ambiente presencial. Nos cursos oferecidos nesta modalidade, a interação entre os envolvidos ocorre através de um Ambiente Virtual de Aprendizagem (AVEA), que possibilita a criação e a (re)utilização de diversos tipos de materiais educacionais em diversos formatos, conhecidos como Objetos de Aprendizagem (OAs). Esta modalidade de ensino já se concretizou como uma prática que ganha cada vez mais adeptos e principalmente uma crescente produção de materiais educacionais. Como resultado, as instituições enfrentam o desafio de ressignificar os recursos que web oferece e de gerenciar este conteúdo elaborado para a modalidade. Neste sentido, a Universidade Federal de Santa Catarina (UFSC) dispõe do Repositório Institucional de Conteúdo Digital (Repositório UFSC), utilizado para o gerenciamento de documentos digitais que vão desde administrativos até outros, de produção científica. Segundo Coelho (2006) um repositório visa à integração do conceito de *Open Access* (Acesso Livre), o que significa a disponibilização livre na



Internet de documentos digitais, permitindo ao utilizador ler, fazer download, copiar, distribuir, imprimir, pesquisar ou referenciar a obra utilizada.

Tendo sua primeira edição com início em 2008, a segunda iniciada em 2011 e a terceira em 2014, o curso de graduação em Letras – Licenciatura em Língua Espanhola e Literaturas de Língua Espanhola, na modalidade a distância, da UFSC, possui atualmente uma estabilidade com demanda periódica e crescente produtividade de materiais criados para as disciplinas disponibilizadas no Moodle-UFSC. Com as novas edições do curso, sentiu-se a necessidade de localizar os OAs utilizados nos semestres anteriores para o planejamento de atividades nas quais pudessem ser reutilizados, independentemente da disciplina. Dessa forma, o Repositório UFSC foi definido como a plataforma para a padronização, organização e o compartilhamento dos OAs, atendendo aos critérios definidos para o projeto em questão, que eram: ser institucional, gratuito e de livre acesso. Assim, o objetivo geral deste trabalho é relatar a metodologia da seleção dos OAs no Moodle e a indexação e usabilidade desses no Repositório UFSC. A fim de alcançar o propósito deste estudo foram trabalhados quatro objetivos específicos: 1) mapear os OAs produzidos para o curso de Letras Espanhol UFSC/EaD; 2) indexar os OAs selecionados no Repositório UFSC; 3) divulgar o repositório aos tutores e professores do curso de Letras Espanhol UFSC/EaD e 4) avaliar o (re)uso da coleção de OAs do curso de Espanhol EaD por parte dos tutores e professores.

FUNDAMENTAÇÃO TEÓRICA

Diante da expansão e intensificação do uso da internet como facilitadora do acesso à informação, é cada vez mais frequente a implementação de repositórios digitais em instituições superiores como forma de gerenciamento da produção intelectual. Sendo uma importante ferramenta para a iniciativa da disseminação e preservação do conhecimento, o repositório possibilita a disponibilização de acervos por meio do movimento de acesso livre (*Open Access*), permite a interoperabilidade entre diversos mecanismos de acesso e armazenamento de dados e dispõe de padrões da iniciativa de arquivos abertos (*Open Archives Initiative – OAI*). Desta forma, os repositórios científicos têm se revelado uma opção importante na disseminação da informação de pesquisa e ensino. (SHINTAKU e MEIRELLES, 2010) Um repositório, de acordo com Santos Junior (2010), pode ser **temático**, restrito a determinado domínio do conhecimento, ou **institucional**, focado na produção de uma instituição específica e composto não só por produção científica, mas também por peças publicitárias, materiais de apoio ao ensino e documentos administrativos. O autor reforça ainda que a utilização de um repositório possibilita não só o gerenciamento de coleções científicas, mas também oferece visibilidade para instituição e diminuição de custos.



Neste contexto, por abranger diferentes tipos de suporte, o Repositório Institucional contribui significativamente para a Educação a Distância, podendo ser utilizado como uma ferramenta de gerenciamento dos OAs produzidos. Sosteric e Hesemeier (2002) concluíram que objetos de aprendizagem equivalem a arquivos digitais compostos por imagens, gráficos, vídeos, sons ou outros recursos educacionais com sugestões de utilização contextualizadas. Os OAs têm como característica principal sua reutilização em diferentes situações de aprendizagem. Rossetto e Moraes (2008) definem ainda que podem ser usados e reusados, adaptando-se facilmente ao longo do processo de aprendizagem de forma a auxiliar os usuários na construção de novos conhecimentos e habilidades. Silva et al (2010) reforçam também a importância da reutilização destes objetos destacando o compartilhamento do recurso como uma meta. As autoras afirmam que com os altos custos de produção do material multimídia essa seria uma medida que permitiria a diminuição de gastos. Sendo assim, Teodoro et al (2008) relembram que as instituições de ensino têm grande preocupação em definir "um padrão de qualidade para a produção de material didático e a sua reutilização, não só como forma de uso racional de recursos, mas também para obter agilidade na oferta de novos cursos."

Criado em 2009, o Repositório Institucional de Conteúdo Digital UFSC teve como objetivo unificar o conteúdo produzido na instituição em um acervo digital central, melhorar a visibilidade da universidade na internet e possibilitar um maior acesso aos conteúdos disponibilizados, dentre outros. Para isso, foi adotado o *Dspace*, software livre, *open source* e direcionado ao acesso aberto. Disponibilizado de forma hierárquica com uma estrutura de comunidades e coleções, o *Dspace* atualmente é amplamente utilizado por diversas instituições, como a melhor forma de organização e gerenciamento da produção científica. O sistema permite a submissão de diversos tipos de materiais digitais, tais como: documentos, livros, teses, publicações multimídia, notícias de jornais, imagens, arquivos de áudio e vídeo, etc. (IBICT...2014)

O curso de graduação em Letras Espanhol EaD/UFSC possui uma grande produção de materiais educacionais para as disciplinas disponibilizadas na plataforma Moodle. Por este motivo, sentiu-se a necessidade de recuperar os OAs produzidos pela equipe pedagógica nos semestres anteriores do curso para o planejamento de atividades e reutilização nas disciplinas das edições atuais. Logo, o Repositório UFSC foi cogitado como a ferramenta mais apropriada para o gerenciamento e compartilhamento dos objetos, visto que este já é utilizada pela universidade.

O referido curso tem duração regular de 9 semestres e, mesmo sendo caracterizado como a distância, acontece de maneira semi-presencial, pois são previstos encontros nos polos de apoio presencial para aulas presenciais, videoconferências, atividades presenciais entre tutores e alunos, atividades de prática de ensino e estágio supervisionado e outras avaliações. Para alcançar os objetivos propostos em seu Projeto Político do Curso, como desenvolver a capacidade de uso da língua estrangeira, nas modalidades oral e escrita, aprimorar o exercício profissional com utilização de tecnologias contemporâneas, utilizar as diferentes linguagens, incentivar o desenvolvimento de projetos multidisciplinares a partir de mídias tais como hipertexto e vídeos, entre outros, o curso de Letras Espanhol EaD utiliza-se de material



didático impresso e digital, disponibilizado por meio do Ambiente Virtual de Ensino e Aprendizagem (AVEA), o Moodle. Os estudantes ainda têm disponíveis como recurso didático um DVD para as disciplinas de língua espanhola, que é entregue aos polos, com exemplos, situações reais, depoimentos de especialistas, esquemas, infográficos, imagens ilustrativas, etc. O AVEA é a base principal para o trabalho do docente e dos tutores e o livro-texto, material de apoio para o professor, contém o conteúdo referente à ementa da disciplina. O livro-texto possui diversos recursos como destaque, link, leitura complementar e outras informações necessárias ao entendimento do conteúdo apresentado. Ainda são elaboradas videoaulas que tratam de assuntos específicos de acordo com as dúvidas dos alunos e necessidade de cada equipe docente.

MATERIAIS E MÉTODOS

Para a realização deste estudo foram necessários diferentes métodos de pesquisa ao longo do processo. Tomando por base o objetivo geral, a pesquisa configura-se como um estudo exploratório, de caráter descritivo. De acordo com Andrade (2001), a pesquisa exploratória caracteriza-se como a fase preliminar, que busca proporcionar mais informações sobre o assunto abordado. Portanto, a elaboração deste estudo, inicialmente, foi realizada com uma revisão de literatura, constituída, em sua maior parte, por artigos científicos voltados a área do EaD e repositórios. Esta revisão serviu como base a um conhecimento maior dos conceitos sobre as temáticas.

Organização dos OAs

Com a finalidade de atingir os dois primeiros objetivos específicos, mapear e indexar os OAs do curso de Letras Espanhol EaD no Repositório UFSC, foram executadas duas etapas, a) triagem do material e b) submissão dos OAs no repositório, conforme descritas a seguir.

a. Triagem do material

O primeiro passo foi definir junto à equipe os critérios para triagem dos materiais contidos no AVEA, tendo como produto final um inventário dos OAs. A partir desta reunião ficou determinada a submissão de materiais explicativos contendo referência, tutoriais, exercícios, videoaulas, videoconferências, imagens e hiperlinks, todos produzidos pela equipe pedagógica do curso de Letras Espanhol EaD. Não foram contemplados os materiais de autoria externa, provas, exercícios direcionados a um único polo, fotos, além de qualquer material produzido pelos alunos.



O inventário criado contém informações básicas que serviram de apoio na segunda etapa do processo de catalogação do material no Repositório da UFSC. Por conta do tamanho dos materiais, os arquivos foram baixados somente no momento da submissão. Pensando em uma maior facilidade ao trabalhar, o inventário foi elaborado na ferramenta *Googledocs*, indicando edição do curso, fase, nome da disciplina, modalidade do material, nome do autor (quando houver), data de inserção no ambiente, além de localização de cada OAs dentro do AVEA, não retirando o link de acesso ao listar o material. Todo o levantamento dos OAs foi realizado desde a primeira edição do curso até a 5ª fase da segunda edição, já que esta encontra-se ainda em andamento.

b. Submissão dos OAs no Repositório

Como o curso de Letras Espanhol EaD foi o primeiro entre os cursos de Licenciatura em Letras da UFSC a disponibilizar os OAs produzidos nas disciplinas do EaD, foi necessário criar uma Comunidade e uma Coleção do curso no Repositório UFSC. Para a criação destas foi preciso enviar uma solicitação à Superintendência de Governança Eletrônica e Tecnologia da Informação e Comunicação da UFSC (SeTIC/UFSC), responsável pelo suporte técnico do Repositório UFSC. Após a solicitação aprovada, a Comunidade e a Coleção foram incluídas no sistema e permitida a delegação a um (ou mais) responsável(is) a inserção de material.

A inserção dos OAs foi realizada pela ordem das disciplinas de Língua Espanhola primeiramente, pois notou-se que estas além de possuírem mais materiais, eram as de maior interesse por parte dos professores e tutores. Em uma segunda conversa com a equipe do curso algumas determinações foram pré-definidas:

- a) os OAs ao serem indexados devem conter uma breve descrição do material;
- b) as palavras-chave devem ser no máximo cinco, sendo três já definidas como “EaD”, “Letras, Espanhol” e nome da disciplina ao qual o material pertence;
- c) caso o OA não possua autor definido no material ficou determinado acrescentar o nome de toda a equipe docente ministrante da disciplina ao qual pertence como autor do objeto;
- d) nas videoaulas e videoconferências deve-se acrescentar além da equipe docente que aparece no vídeo, os nomes da equipe técnica, caso apareçam ao final do vídeo;
- e) para uma melhor recuperação dos objetos, alguns OAs tiveram outras palavras-chave determinadas pelo grupo, fora as três vistas anteriormente, como por exemplo, o material de tutoria ou orientação deve ter a palavra-chave “Tutorial”, ou os exercícios que devem utilizar a terminologia “Atividade didática”;
- f) nas palavras-chave deve ser incluído (além das pré-estabelecidas) somente uma palavra referente ao seu conteúdo.



Divulgação do Repositório

No intuito de atender ao terceiro objetivo específico, optou-se em promover uma apresentação em dois momentos, na Formação de Tutores, em dezembro de 2013, e na Capacitação dos Tutores, em março de 2014. Estes dois eventos já fazem parte do calendário do curso desde a primeira edição. São encontros para novos e antigos tutores, nos quais, junto aos professores e demais envolvidos no curso trocam experiências, discutem acerca dos trabalhos realizados no semestre anterior e sobretudo compreendem e discutem frequentemente o funcionamento do curso na modalidade a distância.

O objetivo de avaliar o (re)uso dos OAs da Coleção do curso de Letras Espanhol EaD por parte dos tutores e professores foi pensado como uma forma de medir a conscientização da usabilidade e o potencial do repositório. A técnica aplicada foi um levantamento de dados através de um questionário on-line (*Googledocs*) enviado via e-mail a todos os tutores e professores do curso de Letras Espanhol EaD e postado no fórum do Espaço Colaborativo da equipe docente do curso no Moodle. O questionário foi elaborado com 11 questões sobre a utilização do repositório e composto de perguntas que auxiliaram a mensurar a proporcionalidade do trabalho feito até o momento, permitindo sua retroalimentação.

RESULTADOS PARCIAIS

Até o momento, o estudo, ainda em andamento, apresenta consideráveis resultados a serem trabalhados e analisados a seguir.

Contando com 41 disciplinas em cada uma das 3 edições, o mapeamento realizado no AVEA totalizou uma média de 1007 OAs, dentre eles, vídeos, hipermídias, textos e imagens. A seleção feita abrangeu toda a 1ª edição e parte da 2ª, já que esta ainda encontra-se em andamento. O mapeamento a partir da 6ª fase e de toda a 3ª edição ainda não foi estabelecido por estarem inseridos em uma versão atualizada do Moodle (versão 2.5).

Em seguida, passou-se à submissão dos OAs no Repositório UFSC. Inicialmente foram tomadas como medidas a conversão de documentos para *pdf* e de vídeos para formato *mp4*, como pedido no regulamento do repositório, porém ao longo do processo notou-se que os vídeos submetidos neste formato falhavam na plataforma e, depois de um tempo foi adotado o formato, *flv*, já utilizado no ambiente Moodle.



As palavras-chave e os resumos foram inseridos conforme os critérios discutidos e aprovados junto à equipe, tornando mais fácil a recuperação da informação. Atualmente já foram indexados 267 OAs, dos quais alguns ao longo do processo de indexação foram reavaliados e descartados, por conta do conteúdo inadequado à proposta, por não possuírem autores ou, ainda, por estarem com a mídia falha ou defasada. Como maior área submetida até o momento estão as disciplinas de língua, seguidas das disciplinas de literatura. Por constituírem um momento de interação entre professor e aluno, as videoaulas e videoconferências são os OAs mais produzidos em todas as disciplinas, conseqüentemente são os objetos mais encontrados na Coleção de Letras Espanhol EaD do repositório.

Com a primeira edição finalizada e duas em andamento, o curso conta atualmente com 49 profissionais na equipe pedagógica, sendo 9 professores, 25 tutores, alocados na UFSC, e 15 tutores polo. Para orientação de toda esta equipe, são realizados encontros quinzenais e semestrais, além de outros periódicos para formação específica tanto de conteúdo quanto de outras habilidades e competências. Como mencionado anteriormente, foram selecionados dois momentos para a divulgação do Repositório. A participação no primeiro evento contou com toda a equipe do semestre (a distância e presencial), já o segundo evento foi destinado unicamente aos tutores e professores que atuam no *Campus* Florianópolis. A divulgação do Repositório UFSC ocorreu na forma de exposição oral, na qual foi conceituado o repositório e exemplificado como realizar buscas de material na plataforma, através de slides com imagens do sistema e permitindo no momento da fala a prática por parte dos ouvintes.

Após as etapas de divulgação, passou-se à etapa de aplicação do questionário. O link do questionário on-line (*Googledocs*) foi enviado para toda a equipe, porém dos 49 profissionais contatados, obtivemos somente 15 respostas. Destes, 9 afirmaram conhecer o

repositório, dos quais, 5 alegaram ter sido pela exposição em um dos dois encontros de equipe do curso. De todos respondentes, apenas 3 afirmaram já ter utilizado a ferramenta e quando o fizeram não obtiveram bons resultados, porém nenhum qualificou a usabilidade do sistema como ruim ou muito ruim. Somente 2 respondentes informaram utilizar a Coleção do curso Letras Espanhol EaD para recuperar material da versão anterior da disciplina em que havia atuado como professor, consultar referências para outras atividades, verificar o que havia sido ministrado nas disciplinas anteriores; relacionar conteúdos, e disponibilizar aos alunos da nova edição materiais elaborados para edições passadas. Algumas questões nem foram respondidas pelos poucos participantes da pesquisa, porém mesmo com um número baixo de respostas já se pode ter um panorama do trabalho desenvolvido até o momento, possibilitando a reflexão de novas estratégias de divulgação do Repositório.



CONSIDERAÇÕES FINAIS

O curso de Letras Espanhol EaD, objetivando o (re)uso dos OAs produzidos em suas disciplinas, foi o pioneiro dos cursos de Licenciatura em línguas estrangeiras da UFSC a elaborar um projeto para criar a sua própria Coleção no Repositório UFSC, visando à organização e disseminação destes materiais à comunidade e demais interessados. Assim, esta pesquisa teve o intuito de relatar a metodologia e a repercussão deste projeto até o momento.

Em relação à etapa de organização, foi possível compreender o quanto é necessário o planejamento prévio na produção do material educacional, pois a maior dificuldade encontrada foi a falta de informação contida nos OAs existentes, como autoria, referência e datas. A falta dessas informações implicou na dificuldade de indexação do material e no descarte de alguns Objetos. Esta problemática nos fez refletir sobre uma estratégia de orientação aos tutores e professores relacionada à importância de registrar este tipo de informação para futura recuperação de conteúdo digital.

Apesar de a submissão de todos os OAs produzidos no repositório ainda não estar concluída, optamos por disponibilizar e divulgar o sistema, já que o curso está com duas edições em andamento e o material poderia auxiliar nas novas disciplinas. Neste quesito, notou-se um interesse por parte da equipe pedagógica em relação à disseminação do conteúdo no repositório, entretanto, consideramos que muitos dos que não acessam é porque não reconhecem na plataforma o seu potencial. Entendemos que é uma mudança de hábito acessar estes recursos e entender como podem ser utilizados para poder realmente apropriar-se dos benefícios que ele oferece.

A partir dessas considerações podemos reavaliar diversos pontos, mas principalmente averiguar a visão e uso dos OAs por parte dos tutores e professores em relação ao trabalho desenvolvido até o momento. Acreditamos que é necessário um trabalho intenso de conscientização em relação à importância de saber pesquisar, acessar e (re)utilizar OAs que estejam em consonância com os objetivos de aprendizagem definidos pelos usuários e que também é fundamental uma nova e constante divulgação do Repositório e da Coleção do curso de Letras Espanhol EaD, desmistificando a imagem estática do recurso e fazendo com que eles possam ressignificar os objetos indexados neste sistema.



REFERÊNCIAS

- ANDRADE, Maria Margarida de. **Introdução à metodologia do trabalho científico**. 5. ed. São Paulo: Atlas, 2001.
- COELHO, Carla. **Um repositório digital para a U.Porto**: relatório preliminar. Porto: Univ. Porto, 2005. Disponível em: <<http://repositorio-aberto.up.pt/bitstream/10216/6701/2/35921.pdf>>. Acesso em: 12 jun. 2014.
- GAZZONI, Alcibiades et al. Proporcionalidade e semelhança: aprendizagem via objetos de aprendizagem. **Revista Renote**, Porto Alegre, 4 n° 2, dez., 2006. Disponível em: <<http://seer.ufrgs.br/index.php/renote/article/view/14141/8076>>. Acesso em: 05 jun. 2014.
- INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. **IBICT**. Disponível em: <<http://www.ibict.br/>>. Acesso em: 10 ago 2014.
- ROSSETTO, Diones; MORAES, Márcia Cristina. Integração entre Repositório de Objetos de Aprendizagem e Ambiente Virtual de Ensino. In: SALÃO DE INICIAÇÃO CIENTÍFICA PUCRS, 9, 2008, Porto Alegre. **Anais...** Porto Alegre: PUCRS, 2008. Disponível em: <http://www.pucrs.br/research/salao/2008-IXSalaoIC/index_files/main_files/trabalhos_sic/ciencias_exatas/ciencia_computacao/62433.pdf>. Acesso em: 05 jun. 2014.
- SANTOS JUNIOR, Ernani Rufino dos. **Repositórios institucionais de acesso livre no Brasil**: estudo delfos. Dissertação (mestrado) – Universidade de Brasília, Faculdade de Ciência da Informação, 2010. Disponível em: <http://repositorio.unb.br/bitstream/10482/5343/6/2010_ErnaniRufinodosSantosJunior.pdf>. Acesso em: 01 jun. 2014.
- SHINTAKU, Milton; MEIRELLES, Rodrigo. **Manual do Dspace**: administração de repositórios. Salvador: EDUFBA, 2010.
- SILVA, Edna Lúcia da; CAFE, Lígia; CATAPAN, Araci Hack. Os objetos educacionais, os metadados e os repositórios na sociedade da informação. **Ci. Inf.**, Brasília, v. 39, n. 3, dez. 2010. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652010000300008&lng=en&nrm=iso>. Acesso em: 28 jun. 2014.
- SOSTERIC, Nike; HESEMEIER, Susan. When is a Learning Object not an Object: a first step towards a theory of learning objects. In: **Internacional Review of Research in Open and Distance Learning**, 2002. Disponível em <<http://www.irrodl.org/content/v3.2/soc-hes.html>>. Acesso em: 05 jun. 2014.



TEODORO, George et al. Proposta para o desenvolvimento de um Repositório de Objetos de Aprendizagem (ROA) na UFMG e UnC. In: CONGRESSO INTERNACIONAL DA ABED , 11, 2008, Santos. **Anais...** Santos: ABED, 2008. p. 1-10. Disponível em: <<http://www.abed.org.br/congresso2008/tc/515200824751PM.pdf>>. Acesso em: 01 jun. 2014.



SISTEMA DE INDICADORES AMBIENTAIS: ARCABOUÇO PARA PROJETOS DE PESQUISA

Kátia V. Cavalcante¹, Claudia A. Tocantins²

¹PhD, Professora Assistente – UFAM kcavalcante@ufam.edu.br

²MSc em Informática, PMP - Fundação Ezute ctocantins@ezute.org.br

RESUMO

Este artigo detalha o processo de implantação de um sistema de indicadores que apoia a sustentabilidade das tomadas de decisão na área de abrangência do estudo socioambiental do Projeto Piatam. O objetivo deste trabalho é propor a metodologia desenvolvida para o projeto de forma a torná-la replicável para outros projetos que construam repositório de dados estruturados para armazenar séries históricas da caracterização ambiental da área de estudo. Uma vez que um protótipo do Sistema de Indicadores Piatam (SIP) foi especificado, projetado e implementado, foi proposta uma metodologia de sistemas de indicadores ambientais passível de ser replicável por outros projetos que envolvam a geração de uma base de dados estruturada para armazenar séries históricas de caracterização de uma área de estudo.

Palavras chaves: Indicadores Ambientais; Sistema de Indicadores, Malária, Amazônia.

ABSTRACT

This paper details the implementation process of a system of indexes supporting environmental management and decision making in the area covered by the Piatam Project. This paper aims to define the methodology developed for the project in order to make it replicable for other projects that hold structured databases (or data repositories) to store historical series of environmental characterization. Once a prototype for the Piatam's System of Indexes (SIP) was specified, designed and implemented, a methodology for environmental index systems was proposed. This methodology is replicable on projects dealing with the creation of structured databases to store historic series of en study area characterization.

Key words: Environmental Indexes; Index System, Malaria, Amazon.



INTRODUÇÃO

Este trabalho apresenta o sistema de indicadores que apoia a gestão ambiental nas tomadas de decisão na área de abrangência do projeto Piatam, um projeto interinstitucional que tem como principal meta a caracterização socioambiental da área de atuação da indústria de petróleo no Estado do Amazonas – Brasil. O objetivo deste trabalho é descrever a metodologia desenvolvida para o projeto de forma a torná-la replicável para outros projetos que construam repositório de dados estruturados para armazenar séries históricas de caracterização ambiental de uma área de estudo. O repositório do projeto insere-se na geração e gestão de conhecimento sobre a ecologia, sociedade e economia da região Amazônica, imprescindível para o desenvolvimento sustentável do país.

As informações armazenadas na Base de Dados Integrada Piatam (BDI) apoiam a criação e disponibilização de um histórico da situação dos recursos naturais na região e podem subsidiar a tomada de decisões em resposta a eventuais impactos ambientais e sociais resultantes das atividades desenvolvidas na área de atuação do projeto. A especificação, o projeto e a implementação da BDI constituem prova de conceito da metodologia empregada.

O Sistema de Indicadores Piatam (SIP), prototipado na última fase do projeto, visou produzir os indicadores do tema Malária de forma automática, para fins de validação e viabilização do sistema a ser desenvolvido. A malária é uma das doenças parasitárias mais importantes entre as que afetam as populações dos países tropicais e subtropicais do mundo. O tema Malária foi escolhido por possuir modelo de dados simples cujo indicador desenvolvido no projeto é composto de dois índices: o índice de atração homem/hora (*IAHH*) e o índice de larva/homem/hora (*ILHH*), diretamente calculados com informações disponíveis na BDI.

A arquitetura do SIP é modular, o que possibilita ser evoluída para atender a novas demandas, de forma a agregar mais valor ao software. É importante ressaltar que os módulos devem atender tanto às demandas de usuários gerenciais, para tomada de decisão, quanto aos usuários técnicos e pesquisadores no ganho de conhecimento. O protótipo possui módulo de geração de gráficos de análise dos dados para auxiliar no diagnóstico da situação ambiental e módulo de geração de layers. O uso de sistemas de informação geográfica (SIG) torna-se uma tendência natural frente a grande diversidade de objetivos com que os dados do SIP podem ser utilizados, pois evidencia, de forma visual, as relações da informação com o seu meio físico. No contexto da disseminação da informação, o que se buscou é aumentar a abrangência dos resultados do Projeto, portanto, além de instrumentalizar, em um contexto espacial, os indicadores, estar-se-ia auxiliando no planejamento e gestão ambiental aplicada à política pública.

HISTÓRICO

Desde sua primeira fase, o Projeto Piatam definiu como resultado esperado a organização de uma base de dados que agregasse e disponibilizasse os dados obtidos nos diversos temas que compunham a multidisciplinariedade da pesquisa. Em Cavalcante, *et.all* (2007), foi apresentado um conjunto de indicadores da dinâmica dos processos biológicos, socioeconômicos e dos meios físicos da Amazônia, no trecho na área de influência do trecho Coari – Manaus.

Os critérios selecionados para estabelecer os indicadores Piatam foram: Existência de dados representativos em base de dados; Possibilidade de cálculo da sua quantificação; Possibilidade de manter a série histórica atualizada; Facilidade de interpretação.

Tais indicadores foram construídos a partir das séries históricas coletadas em campo e analisadas em laboratório, residentes na BDI. Uma continuação natural da pesquisa foi apresentada no âmbito do Grupo de Tecnologia da Informação com o objetivo de investigar formas de automatizar o cálculo dos indicadores propostos em cada área, para extrair informações e apoiar as análises dos dados na continuidade das coletas de campo.

PROTÓTIPO DO SIP

Objetivo

O trabalho desenvolvido compreendeu a especificação, projeto e implantação de um sistema preliminar para a prova de conceito do SIP, que apresentou indicadores selecionados e calculados, a partir de dados oriundos da BDI. Os números e gráficos extraídos automaticamente da BDI seriam, então, comparados com dados pretéritos publicados para efeito de avaliação da exequibilidade do cálculo contínuo dos indicadores à medida que a coleta de dados evoluísse.

Metodologia Aplicada

A especificação do SIP incluiu as seguintes atividades:

(1) Análise da Proposta do SIP - O primeiro passo da metodologia foi o estudo da proposta sugerida em (Ibid, 2007) que buscou entender, a construção de indicadores e atributos que possibilitam análises mais complexas e completas das realidades social e ambiental da área de estudo. Considerando que os protocolos metodológicos de cada área temática estão consolidados no modelo de dados, na BDI Piatam, realizou-se uma pré-seleção que serviu para orientar e ajustar os métodos de construção dos indicadores, ocorrendo, assim, a seleção e a validação final destes.

(2) Projeto de Arquitetura do SIP - A arquitetura proposta para o SIP é modular para permitir portabilidade, escalabilidade e segurança do sistema, bem como a inclusão de novas funcionalidades, a saber: *Business Intelligence* e *Data Mining* (Mineração de Dados) com o intuito de agregar mais valor ao produto.

(3) Análise das fórmulas de cálculo versus Modelo de dados da BDI, para o tema escolhido - Os indicadores foram calculados com base nas fórmulas dos índices *IAHH* e *ILHH*. Realizou-se a simulação dos cálculos, a partir dos dados armazenados na BDI Piatam, o que apoiou a análise e o projeto de algoritmo da implantação do SIP; Foram, ainda, elaborados diferentes gráficos (barra, linha, setores e rosca) de indicadores entomológicos do tema malária (*IAHH* e *ILHH*) no Microsoft Excel, a partir de dados extraídos da base. Estes gráficos foram analisados para atender tanto às demandas de usuários de nível gerencial, para tomada de decisão (gráficos gerenciais), quanto aos usuários técnicos (pesquisadores).

(4) Implementação de versão preliminar do SIP na BDI - As seguintes atividades foram executadas nessa fase: elaboração do documento de especificação de casos de uso para os desenvolvedores; elaboração da tela de filtro do protótipo; criação de *queries* para obtenção de dados da BDI; criação de *views* em substituição às *queries* para facilitar a implementação e otimizar o processamento; codificação; e integração.

(5) Estudo para a Visualização Georreferenciada dos Indicadores - Estudo para o desenvolvimento do módulo de geração de *layers* que apresenta os indicadores ambientais entomológicos georreferenciados, pontuando e quantificando as áreas de incidência da malária no trecho de atuação do Projeto.

Arquitetura do SIP

O diagrama de blocos da arquitetura do SIP, ilustrado na Figura 01, retrata os principais subsistemas e suas interconexões. Essa arquitetura, definida em três camadas (interface, controle e negócio), utiliza componentes de software de tecnologia J2EE¹ para fornecer portabilidade, escalabilidade e segurança do sistema. Esses componentes fazem parte da arquitetura geral do aplicativo de consulta da BDI discutidos na referência (BORELLI *et al*, 2005). As funcionalidades dos subsistemas são descritas a seguir:

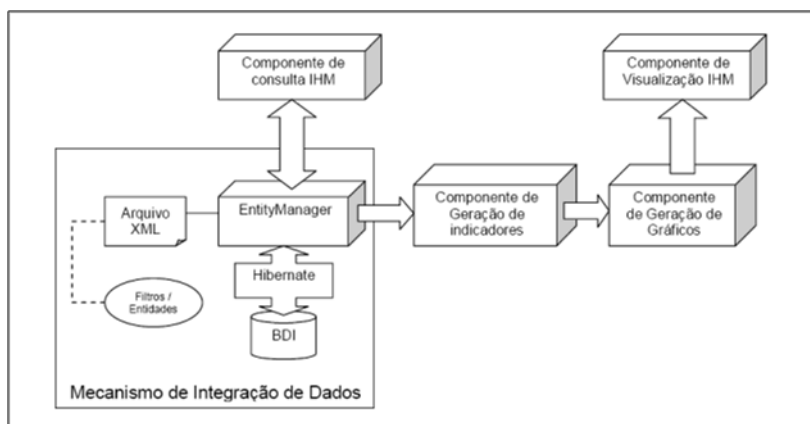


Figura 01. Diagrama de Blocos do SIP

¹ J2EE – Java 2 Enterprise Edition é uma plataforma de programação para servidores na linguagem de programação Java.

– A componente de consulta interface homem/máquina – IHM é a interface gráfica de usuário para a parametrização de indicadores. Ela permite que sejam definidos pelo usuário, filtros de busca comuns inerentes ao modelo, organizados, de forma intuitiva, para a determinação dos dados a serem retornados e os critérios de seleção no tempo e no espaço.

– *Entity Manager*² – Gerenciador das entidades do sistema capaz de integrar os vários dados, interfaceando as definições e filtros com o mecanismo de persistência. Este componente permite criar um mecanismo de consulta de dados eficiente e diversificado.

– *Hibernate*³ – Software livre para mecanismo de persistência de objetos. Faz parte do mecanismo de integração de dados da base de dados.

– BDI – Sistema Gerenciador de Base de Dados (hoje, o Oracle 9g). Este componente faz parte do mecanismo de integração de dados da BDI Piatam.

– Componente de Geração de Indicadores – faz o cálculo dos indicadores, a partir dos dados de parâmetros da consulta.

– Componente de Geração de Gráficos – de posse dos dados da consulta e indicadores calculados no subsistema anterior, geram-se os gráficos para análise dos dados: gráficos de barra, linha e de setores.

– Componente de Visualização de IHM – interface gráfica de usuário para visualização dos gráficos gerados pelo subsistema anterior.

MODELO DE DADOS – MALÁRIA

Modelo Lógico

O modelo lógico do tema Malária, ilustrado na Figura 02, apresenta o relacionamento entre as principais entidades do modelo, incluindo, em destaque, as entidades do cabeçalho. O cabeçalho é um componente espaço-temporal utilizado para integrar os dados das áreas temáticas da BDI. Toda amostra de dados coletada no Projeto é relacionada a um ponto de coleta, que é definido pela associação de componentes espacial e temporal.

² O termo em inglês *Entity Manager* refere-se ao uso de um padrão de projeto de programação (*Design Pattern*) largamente utilizado em práticas de projeto orientado a objetos e, neste caso, refere-se à classe que gerencia os objetos do tipo “Entidade”.

³ *Hibernate* é um framework para o mapeamento objeto-relacional escrito em linguagem Java. Ele facilita o mapeamento dos atributos entre uma base de dados relacional e o modelo de objetos de uma aplicação, mediante a definição em arquivos XML.

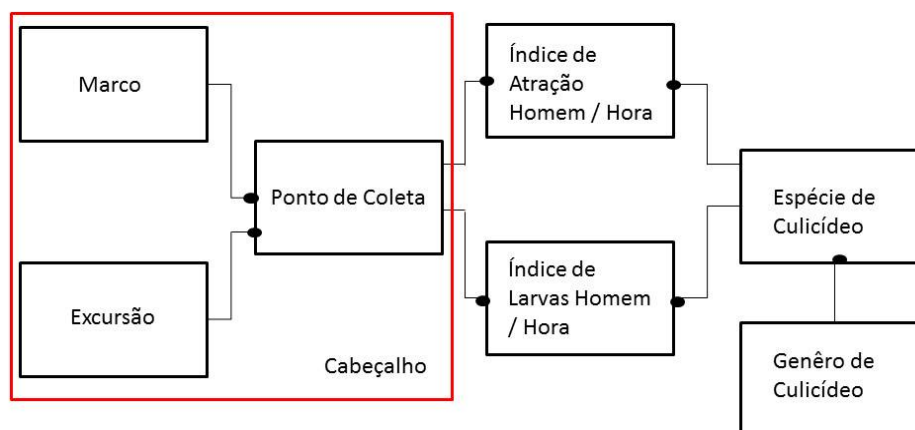


Figura 02. Modelo lógico BDI - Tema Malária

No Piatam, a componente espacial é definida pelo Marco. A componente temporal é definida pelas excursões de coleta de campo. Esse componente foi desenvolvido com a intenção de integrar dados de natureza diversa, permitindo a análise conjunta de diferentes áreas temáticas da BDI (TOCANTINS, *et.all.* 2005).

A partir de um ponto de coleta (associado à excursão e marco do cabeçalho), se obtém o Índice de Atração por Homem/Hora (*IAHH*), mediante parâmetros obtidos no local de coleta, como: número de exemplares coletados; número de coletores; número de horas de coleta; e número de locais de coleta. Associada ao *IAHH*, está a descrição de espécie de culicídeo (mosquito) para caracterização das espécies coletadas. Da mesma forma, associada à informação da espécie está a descrição do gênero de culicídeo para a caracterização genérica das espécies. Para as informações de Índice de Larva por Homem/Hora (*ILHH*), usam-se os mesmos parâmetros do *IAHH* e a mesma caracterização de espécies e gêneros.

Modelo físico

O modelo físico do tema Malária está ilustrado na Figura 03. Nesse modelo, os campos em destaque são utilizados nos cálculos dos indicadores e na geração de gráficos para análise dos dados.

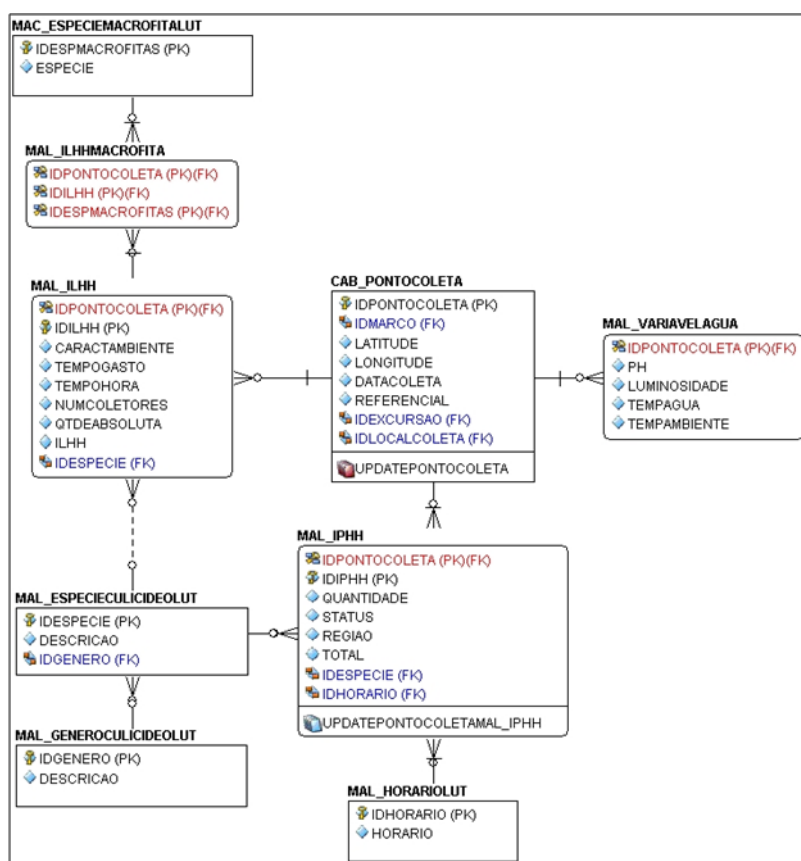


Figura 03. Modelo Físico - Tema Malária

Cálculo dos Indicadores Entomológicos

Índice de Atração Homem/Hora - IAHH

O *IAHH* é resultante de uma relação entre o número de mosquitos capturados, o número de coletores, o tempo de coleta e o número de localidades estudadas. O índice é calculado com base na equação 1 (CAVALCANTE, et all, 2007).

$$IAHH = \sum_{i=1}^L \left(\frac{N}{C \times h} \right) \quad (1)$$

O parâmetro *N*, da equação 1, se refere ao atributo quantidade, número de exemplares de mosquitos capturados. O atributo horário, por sua vez, é a string "18-22h" (conteúdo-padrão para todos os registros), que precisa ser convertido para número de horas *h*, a fim de ser usado na referida equação. Para validação do protótipo, foram considerados, respectivamente, os valores 1 (apenas um coletor) e 4 (interpretado de 18-22h) aos parâmetros *C* e *h* da equação 2. Considerando o padrão da série histórica armazenada.

$$IAHH = \left(\frac{N}{C \times h} \right) \quad (2)$$

Como exemplo, o cálculo do índice *IAHH*, para o marco Santa Luzia do Baixo (Coari-AM), no ano de 2004, tem-se:

$$IAHH = N / (C \times h) = 21,4 / (1 \times 4) = 5,36.$$

Para o marco de Santo Antônio (Anori-AM), no mesmo ano, calcula-se:

$$IAHH = (N / C \times h) = 95,64 / (1 \times 4) = 23,91.$$

Índice de Larvas por Homem-Hora - ILHH

O *ILHH* resulta de uma relação entre: o número de larvas capturadas, dividido pelo número de coletores e pelo número de horas de coleta e o número de locais de coletas. O índice é calculado com base na equação 3.

$$ILHH = \sum_{l=1}^L \left(\frac{N}{C \times h} \right) \quad (3)$$

O atributo *Tempo/Hora* é igual ao número de horas de coleta *h*, o atributo *NumColetores* é igual ao número de coletores *C*, o atributo *QtdeAbsoluta* é igual ao número de exemplares de larvas capturadas *N*. Além disso, este modelo possui o atributo *ILHH*, que é o resultado da aplicação desses parâmetros na equação 3 (assumindo *L* igual a 1). Sendo assim, este índice pode ser redefinido pela equação 4.

$$ILHH = \left(\frac{N}{C \times h} \right) \quad (4)$$

Como exemplo do cálculo do índice *ILHH*, para o marco Bom Jesus (Manacapuru-AM), no ano de 2004, tem-se:

$$ILHH = N / (C \times h) = 53,04 / (1 \times 4) = 13,26.$$

Para o marco Matrinxã (Codajás-AM) no mesmo ano, tem-se:

$$ILHH = N / (C \times h) = 83,44 / (1 \times 4) = 20,86.$$

APRESENTAÇÃO DOS INDICADORES

Os gráficos são instrumentos que facilitam a compreensão de comparações de padrões e tendências nos dados, mediante a sua interpretação visual, e são utilizados para analisar o comportamento dos índices entomológicos, por comunidade ou por período hidrológico. São utilizados gráficos de barra, linha e rosca para destacar os dados dos indicadores oriundos da BDI de vários modos.

- O gráfico de barra mostra as alterações de dados em um período de tempo ou ilustra comparações entre itens.
- O gráfico de linha mostra tendências nos dados em intervalos iguais.
- O gráfico de rosca, por sua vez, similar ao gráfico de setores, apresenta o relacionamento das partes com o todo, tornando-se útil quando se deseja dar ênfase a um elemento importante.

Os dados dos indicadores, com vistas à elaboração dos gráficos desta seção, foram gerados, a partir de consultas SQL⁴ da BDI de produção referente aos anos de 2004 e 2006.

Gráficos para IAHH anual por comunidade

Os gráficos de barras e rosca representam, de forma mais legível, o comportamento anual de índices IAHH por comunidade.

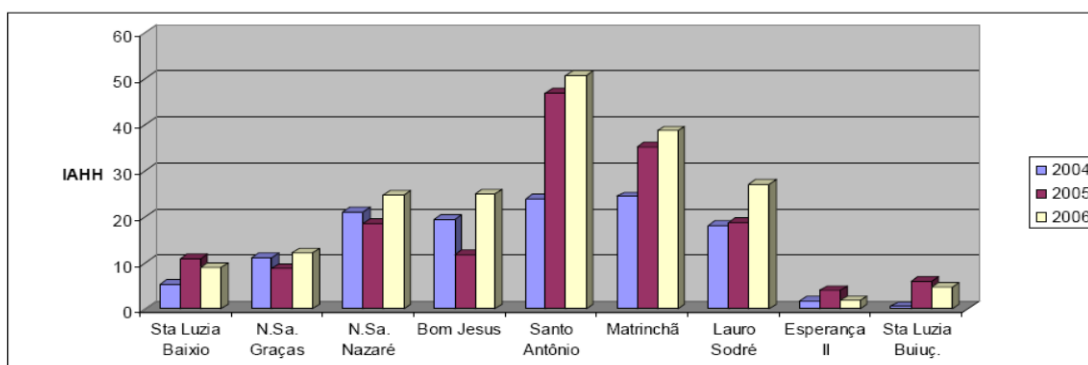


Figura 04 Todos os mosquitos coletados nos anos de 2004 a 2006, por comunidade.

Gráficos para ILHH por Período Hidrológico

O gráfico de linha representa, de forma mais legível, a tendência de índices ILHH, por período hidrológico anual. Nas Figuras 05 e 06, é ilustrada esta tendência para todas as larvas coletadas, por período hidrológico, no ano de 2004 e 2006.

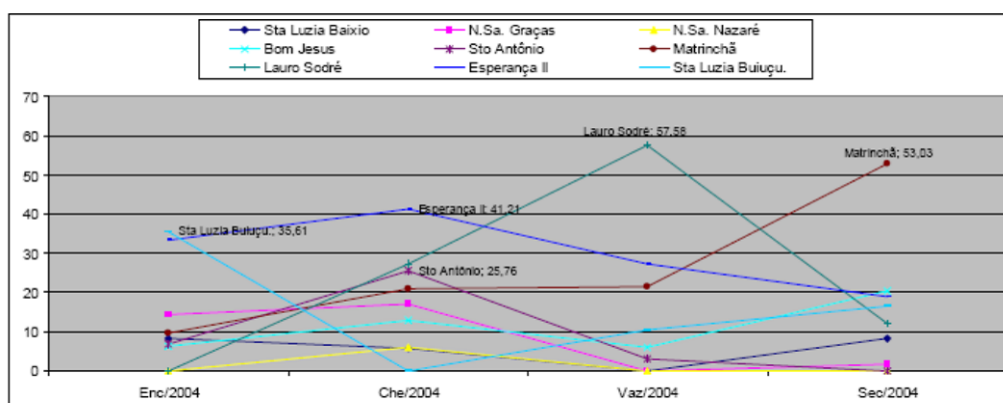


Figura 05. O ILHH, por período hidrológico, no ano de 2004

⁴ Structured Query Language ou Linguagem de Consulta Estruturada, linguagem de pesquisa declarativa padrão para banco de dados relacional.

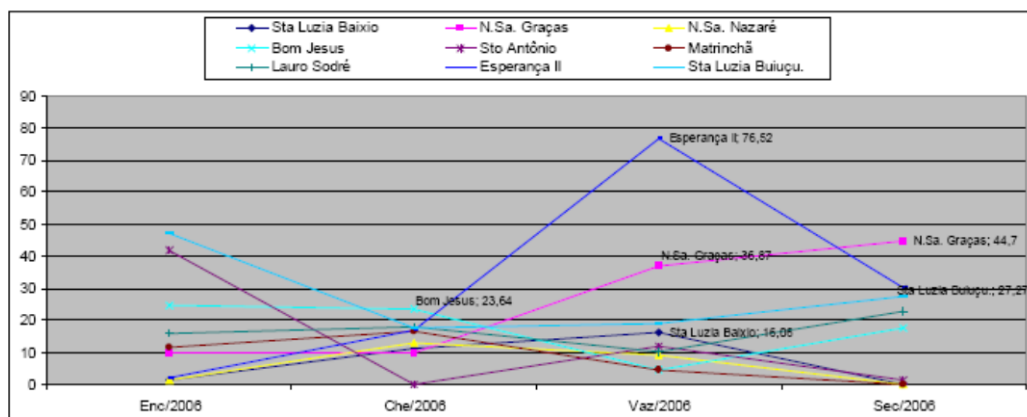


Figura 06. O *ILHH*, por período hidrológico, no ano de 2006

Implementação da versão preliminar do SIP

A partir da definição dos gráficos, o sistema foi implementado, em Java, dentro do aplicativo de gestão da BDI Piatam⁵. a definição de parâmetros para o cálculo do indicador, sendo possível o uso dos seguintes filtros: tipo de indicador, espécie, comunidades e anos; o sistema emite os relatórios.

Estudo para a Visualização Georreferenciada dos Indicadores

Após a implantação da versão preliminar do SIP, avaliou-se a necessidade da visualização georreferenciada. A sequência de atividades, utilizada para o desenvolvimento do módulo de visualização, previu três etapas, assim descritas: organização do banco de dados georreferenciado, construção da base cartográfica, e extração das informações que satisfaçam as formas de apresentação e facilitem a interface com o usuário.

Para construção da base cartográfica, os procedimentos foram descritos por meio de fluxo esquemático capazes de fornecer suporte à análise integrada do ambiente apresentado. O fluxo envolveu a geração de um mapa de áreas sujeitas a inundação, a partir dos dados de hidrografia, solos, altimetria e declividades. Este foi combinado com um mapa de cobertura vegetal extraído da cartografia “uso da terra” para identificar áreas de potencial criadouros. O potencial de ocorrência de imaturos é calculado pela combinação do potencial de criadores gerado mais restrições dadas pelo PH e luminosidade e a presença de macrófitas. Esse mapa é associado à proximidade de ocorrência de malária, baseado na notificação de casos e com o potencial de ocorrência de alados (construído por sua vez pela proximidade dos bolsões x proximidade de criadores de animais x presença de núcleos urbanos x permeabilidade da paisagem). O mapa resultante localiza o Risco de Proliferação da Malária, validado pelos valores obtidos em coleta de campo, gerando os mapas

⁵ O aplicativo está descrito em (Cavalcante, Tocantins 2007).

finais de IAHH e ILHH para a espécie do Gênero *Anopheles* coletadas por comunidades.

A preparação dos planos de informação para o processo criação dos *layers*, visa a identificar e a definir as áreas com potencial de vetorização. Foram construídas, separadamente, *layers* com os pontos georreferenciados, para o IAHH e para o ILHH. Em razão da grande densidade da malha hidrográfica, trabalhou-se, então, com lagos (principais pontos de coleta do ILHH) e canais de acesso aos lagos e ao rio (igarapés, paranás). Em seguida, foram criados buffers 500m, 1 km, 3 km e 5 km de distância, para todas as linhas desse *layer*, criando, assim, uma hidrografia de polígonos. Esta distância foi estimada por meio da média de desempenho de vôo do mosquito transmissor da malária. O mosquito da malária pode voar por 1 a 4 horas continuamente até 1-2 km / h viajando até 10 km em uma noite. (KAUFMANN & BRIEGEL, 2004).

Foram então gerados *layer* de hidrografia para o período da seca e período da cheia, buscando-se caracterizar e relacionar a sazonalidade epidemiológica da malária, com identificação dos corpos de água, uma vez que nos meses de chuvas intensas (janeiro a abril), estas funcionam como um dos controladores naturais da densidade *Anopheles darlingi*. Entre abril e julho com a diminuição das chuvas aumentam os criadouros apropriados a proliferação deste anofelino. (TADEI, 1996)

As operações finais dos *layer* intermediários deram-se a partir do cruzamento dos seguintes temas: *Índice de Atração Homem-Hora* = Uso do Solo x Cobertura Vegetal; *Índice de Larvas Homem-Hora* = Hidrografia (seca ou cheia) x Cobertura Vegetal. *Índice de Larvas Homem-Hora* = Hidrografia (seca ou cheia) x Tipologia da Água branca e Água Preta

O procedimento de combinação para o ILHH é descrito pela fórmula abaixo⁶. O mesmo procedimento é realizado para o IAHH.

$$A_{ij} = (TC.BD) C^n$$

Onde:

A_{ij} = Local do *layer*

TC = Tipo de cobertura vegetal/hidrologia (nota)

BD = peso obtido no *buffer* de distância (peso)

C = Tipo de Criadouro

$$A_{ij} = (TC. BD) C1 + A_{ij} = (TC. BD) C2 + A_{ij} = (TC. BD) C3 + A_{ij} = (TC. BD) C4 + A_{ij} = (TC. BD) C5$$

O segundo conjunto de *layer's* construído com os dados alfanuméricos advindos da base de dados destacava a quantidade capturada tanto de imaturos quanto de alados para o gênero *Anopheles*. Esses *layers* foram denominados de Mapa: IAHH Anual para a Espécie do Gênero *Anopheles* Coletadas por Comunidade; e Mapa: ILHH Anual para a Espécie do Gênero *Anopheles* Coletadas por Comunidade. Com sobreposição das *layers* na sequência acima descrita foi gerado o Mapa final: Dinâmica da distribuição de ocorrência de imaturos (ILHH) e alados (IAHH) no trecho de atuação do Piatam.

Finalizada a fase operacional de preparo da base de dados espaciais e do banco de dados onde foram armazenados os atributos, iniciou-se a fase de definição dos procedimentos para extração das informações que satisfaçam as formas de apresentação, obedecendo a critérios de representação cartográfica.

⁶ Adaptado do procedimento de XAVIER da SILVA, [11].



CONSIDERAÇÕES FINAIS

A natureza investigativa dos processos de pesquisa impõe um desafio grande para a continuidade de séries históricas. Descobertas científicas e a própria evolução das pesquisas resultam em alteração nas metodologias de coleta e nos modelos de interpretação dos dados. A eficácia dos indicadores, porém, dependem da sua capacidade de comparar a área estudada ao longo do tempo.

Os resultados obtidos até o momento e a experiência acumulada ao longo da evolução da base de dados indicam que a aplicação desta metodologia em outros projetos multidisciplinares é viável e interessante, e contribui, de forma objetiva, para a avaliação da sustentabilidade em projetos de pesquisa, por atender ao desafio da construção de um sistema indicadores para apoio à gestão ambiental sustentável, característica comum a maior parte dos projetos da área socioambiental.

A condição para a metodologia proposta seja utilizada é que exista a definição de indicador(es) que sigam os critérios apresentados na seção 2 e uma base de dados cujo modelo é estruturado e preparado para receber os parâmetros utilizados no cálculo do indicador. A viabilidade final do projeto é dada pela disponibilidade dos dados em si na base de dados, ou seja, que haja continuidade temporal e metodológica na coleta de campo dos dados presentes nas respectivas fórmulas de cálculo.

Agradecimentos

As autoras são gratas à Financiadora de Estudos e Projetos, FINEP, pelo apoio financeiro por meio do Projeto Piatam.

As autoras são gratas à Fundação de Amparo a Pesquisa do Estado do Amazonas – FAPEAM, pelo apoio financeiro que viabilizou a nossa presença neste renomado evento.

Agradecem à coordenação geral do projeto e as inúmeras contribuições da equipe no desenvolvimento da BDI.

REFERÊNCIA

BORELLI, FREDERICO, TOCANTINS, CLAUDIA A., FURTADO, PHILLIP. *Um Mecanismo Baseado em XML para Integração de Dados Heterogêneos na BDI-PIATAM/SIPAM.* In Anais do I Congresso Internacional do PIATAM: ambiente, homem, gás e petróleo. Manaus, Brasil, 2005.

CAVALCANTE, Katia V., RIVAS A.A.; FREITAS C.E. *Indicadores Socioambientais e Atributos de Referência para o trecho Urucu-Coari-Manaus,* Rio Solimões, Amazônia Ocidental, EDUA, 2007.

CAVALCANTE, Katia V.; TOCANTINS, CLAUDIA. A. *FBDI PIATAM: Uma ferramenta para a integração e disseminação de informações em projeto multidisciplinar.* In: SIGE 2007, 2007, Santarém. Consolidando a Ciência da Informação na Amazônia - SIGE/Anais. Santarém, 2007.



JUNK, W.J e FURCH, K The physical and chemical properties of Amazonian waters and their relationships with the biota. In: Key Enviroments Amazoniana. Eds.: G.T. Prance e T.E. Lovejov. Pergamon Press.1985

KAUFMANN, Christian; BRIEGEL, Hans. Flight performance of the malaria vectors *Anopheles gambiae* and *Anopheles atroparvus*. Journal of Vector Ecology. June, 2004.

MOTTA, R. S. *Indicadores ambientais no Brasil: aspectos ecológicos, de eficiência e distributivos.* Brasília: IPEA, 1996. 104p. (Textos para discussão, 403).

MOURA, Ana Clara Mourão *Tendência recente nos estudos urbanos e o papel da Cartografia Temática.* Belo Horizonte: Caderno de Arquitetura e Urbanismo, 1993. pág. 23-25.

TADEI, W. P. Projeto: Malária – Fatores Entomológicos envolvidos na Dinâmica de Transmissão. I. Densidade populacional, Caracterização dos Criadouros e Sazonalidade. Relatório Final/PCMAM. MCT, INPA, 1996.

TOCANTINS, C A., EVANGELISTA FILHA, I. M., CAVALCANTE, K. V. *Dados Multidisciplinares: integração espaço-temporal.* In: Anais do 1º. Congresso Internacional PIATAM: ambiente, homem, gás e petróleo. Modelagem ambiental. Pág.160, 2005.

XAVIER da SILVA, J Geoprocessamento e Análise Ambiental. **Revista Brasileira de Geografia.** 54 930: 47-61, 1992

CHAIN-REDS DART Challenge

*Roberto Barbera^{a,b}, Bruce Becker^c, Carla Carrubba^b, Giuseppina Inserra^b,
Salma Jalife Villalón^d, Christos Kanellopoulos^e, Kostas Koumantaros^e, Rafael
Mayo-García^f, Luis A. Núñez^g, Ognjen Prnjat^e, Rita Ricceri^b, Manuel Rodriguez
Pascual^f, Antonio Rubio-Montero^f, Federico Ruggieri^{h,i} (On behalf of the
CHAIN-REDS project)*

^a Dpt. of Physics and Astronomy, University of Catania, Viale A. Doria 6, Catania 95125, Italy

^b National Institute of Nuclear Physics-Catania, Via S. Sofia 64, Catania 95123, Italy
{roberto.barbera, carla.carrubba, giuseppina.inserra, rita.ricceri}@ct.infn.it

^c Meraka Institute, Meiring Naudé Road, Pretoria, 0001, South Africa bbecker@csir.co.za

^d Corporación Universitaria para el Desarrollo de Internet, Parral No. 32 Col. Condesa 06140
Mexico D.F., Mexico salmajalife@cudi.edu.mx

^e Greek Research and Technology Network, 56 Mesogion Av., Athens, 11527, Greece {skanct,
kkoum, oprnjat}@admin.grnet.gr

^f Centro de Investigaciones Energéticas Medioambientales y Tecnológicas, Av. Complutense 40,
Madrid, 28040, Spain {rafael.mayo, manuel.rodriguez, antonio.rubio}@ciemat.es

^g Escuela de Física, Universidad Industrial de Santander, Cra 27 calle 9, Bucaramanga, Colombia
lnunez@uis.edu.co

^h Consortium GARR, Via dei Tizii 6, Rome, 00185, Italy

ⁱ National Institute of Nuclear Physics-Roma Tre, Via Vasca Navale 84, Rome, 00146, Italy
federico.ruggieri@roma3.infn.it

ABSTRACT

CHAIN-REDS (Coordination and Harmonisation of Advanced e-infrastructure for Research and Education Data Sharing) is EU project focused on promoting and supporting technological and scientific collaboration across different communities established in various continents. Nowadays, one of the most challenging scenarios scientist and scientific communities are facing is huge amount of data emerging from vast networks of sensors and from computational simulations performed in a diversity of computing architectures and e-infrastructure. The new knowledge coming out from the interpretation of these datasets, reported on the scholar literature, is increasingly problematic to be reproducible due to the difficulty to access measured data repositories and/or computational applications that generate synthetic data through computer simulations. This paper presents CHAIN REDS approach, several tools and services, based on the adoption of standards, aimed at providing easy/seamless access to datasets, data repositories, open access document repositories and to the applications that could make use of them. All these tools and services are enclosed in what we



have called the Data Accessibility, Reproducibility and Trustworthiness (DART) challenge. This initiative allows researchers to easily find data of his interest and directly use them in a code running by means of a Science Gateway (SG) that provides access to cluster, Grid and Cloud infrastructure worldwide. In this scenario, the datasets are found by means of either the CHAIN-REDS Knowledge Base (KB) or the Semantic Search Engine (SSE), the applications ran on the CHAIN-REDS SG, accessible through an Identity Federation. The datasets can be both identified by Persistent Identifier (PID) and assigned unique number ID. Scientists can then access the data and the corresponding application in order to either reproduce and extend the results of a given study or start a new investigation. The new data (and the new paper if any) are stored on the Data Infrastructure and can be easily found by the people belonging to the same domain making possible to start the cycle again.

Keywords: Global research; Open access; Data management; Standards; Identity provision; e- Infrastructure.

INTRODUCCIÓN

Data-intensive scientific analysis is a completely new way of doing science. How to deal with large datasets is still in evolution and has a long way to go. All disciplines, either physical, life sciences and humanities are becoming increasingly data-driven and data intensive. This is happening mainly due to technological advances in information networks, computing capacity, big instruments, penetration by sensors in all areas, as well as increasingly collaborations among researchers. Big Data Science requires interdisciplinary skills in which computer scientists, statisticians and other experts combine their knowledge to create new techniques, tools and methodologies, shifting from a hypothesis-driven to a data-driven way to analyse the increasing datasets. Astronomy is the most pioneer data-driven science and its communities are early adopters and creators of multiple discovery environment incorporating strategies and tools to manipulate and analyse huge amounts of data.

Large data productions are usually carried out by global collaborations, i.e., multinational science groups that generate large volumes of data, geographically distributed and maintained only during the project life cycle. Most of these data is never published and, when the collaborations end, many is lost or stashed away in national (or international) reservoirs that have nothing to do with their origins. Production decisions, approximations and provenance are buried in a huge electronic correspondence to which no-one has access.



A similar path is followed by small data producers scattered around the globe. Both large and small data producers face the same problems in knowledge cataloguing, preservation and dissemination. It is imperative to plan and build repositories that store data as they emerge and to retain the history of the decisions and criteria that generate them. Starting the century several multilateral organizations and planners in Europe and the United States generated technical reports to encourage the preservation of important scientific data collections. Recently, most of these recommendations have rooted as national and multinational initiatives for general policies concerning data curation. However, many of these recommendations have not permeated into the producing communities and/or to the collection custodians in these countries. The situation is even worse in Latin America where we are still not convinced by, or at least aware of, the new paradigms in the production and dissemination of scientific knowledge, and consequently, only a low-level use of Information and Communication Technology (ICT) awareness has been incorporated in teaching and research.

In this emerging context CHAIN-REDS (Coordination and Harmonisation of Advanced e-infrastructure for Research and Education Data Sharing) is EU project focused on promoting and supporting technological and scientific collaboration across different communities established in various continents. To do so, it is essential to promote instruments and practices that can facilitate their inclusion in the community of users, i.e. the use of standards. Then, to build on the best practices currently adopted in Europe and other continents, and promote and facilitate interoperability among different e-Infrastructures is a must. CHAIN-REDS, in accordance with several European strategies, plans to focus on including low-level services, exchanging in data infrastructures and support preservation and data exploitation services, as well as activities aimed at interoperability and data access federation and openness. Addressing basic issues such as data persistency, accessibility and interoperability will be the first general goal.

As a main issue, the efficient access, use and further analysis of Data has emerged. The number of Data Repositories (DRs), either Open access ones (OADRs) or not, and the quantity of TB they store have largely increased in the latest years. As a consequence, if CHAIN-REDS aims to allow VRCs, research groups and single researchers to efficiently use worldwide distributed resources, it is needed that the data they are employing will be interoperable as well. Otherwise, advances made on middleware interoperability will result meaningless since the computational resources will not be properly exploited.

This paper we shall describe the ecosystem of the main CHAIN REDS data tools and how this ecosystem of tools and services are used to help solving what we have called the Data Accessibility, Reproducibility and Trustworthiness (DART) challenge.

INTEGRATION OF DATA RELATED CAPABILITIES

In the current days, there have been extraordinary advances in the network and computational capacities. Just to mention a few correlated ones, academic networks have made available distributed infrastructure as grid or cloud and Infiniband links have deeply increased the parallel performances. Nevertheless, a major new challenge has arisen due to the huge amount of computational calculations and services that have been made: the management of data. Both in academic and scientific fields, the stored data have dramatically increased and, even more, their use is demanded by more and more people. This fact can be easily showed: the number of Data Repositories (DRs) and Open Access Document Repositories (OADRs) and the volume of data they store have largely increased in the latest years. As a consequence, it is necessary that the data will be easily used as well; furthermore, when these data concern datasets and publications. Otherwise, advances made on middleware interoperability will be meaningless since the computational resources will not be properly exploited. In this regard, CHAIN-REDS is promoting interoperability as a main objective and a worldwide demo has been recently shown in September 2013 covering different regions worldwide.

To achieve DART, several tools have been implemented by the project. The Knowledge Base (KB) provides information about the deployment of e-Infrastructure related topics per country and even about specific Distributed Computing Infrastructures (DCIs) by means of a Site or a Table view. During 2013, the project has been working on extending the CHAIN KB with information related to data infrastructure. To do so, it has collected both issues and best practices and has surveyed the involved regions in order to discover data repositories. The reason for that is to promote data sharing across different e-Infrastructure and continents widening the scope of the existing CHAIN KB to Data Infrastructure and to finally provide proof-of principle use-cases for data sharing across the continents.

Before describing the CHAIN-REDS tools we shall present in the next section, two important initiatives strategies closely related to link data to other type of digital contents.

PROMOTING IDENTITY PROVISION AND DATA STANDARDS

One of the major challenges in providing services is how the users access them, i.e. how they authenticate themselves and which roles are allowed to assume over that services. For many years, databases with information on users (username and password) have been provided by the service managers and more restricted solutions have been taken also into account such as that implemented to access Grid computing (personal certificates provided by an accredited Certification Authority). Nevertheless, those solutions have usually driven the users to a wide set of usernames and password pairs (with the difficulty of remembering all of them). In the latest years, the concept of Identity provision has emerged as a valid solution. Furthermore, such a concept is of importance in the



Academia, where every student, professor or administrative staff has his/her username and password as he/she becomes part of an institution. Thus, being a University, an R&D Centre or an NREN accredited by an Identity Provider, a huge pool of services can be accessed by a single user with only a pair of associated credentials (username and password).

Such access has been successfully demonstrated in Grid computing, where no more personal certificates are needed and where now robot certificates can manage jobs for a long period of time, but it can be also applied to academic services for students such as those closely related to a Faculty Secretariat or to their academic record just to mention a few.

In addition to foster identity provision CHAIN-REDS has selected to promote the following standards for pursuing DART initiative:

- OAI-PMH⁶ for metadata retrieval.
- Dublin Core⁷ as metadata schema.
- SPARQL⁸ for semantic web search.
- XML⁹ as potential standard for the interchange of data represented as a set of tables.

To those, Persistent Identifiers (PID)¹⁰ must be added as a tool to know where and how data and metadata are stored. Such a circumstance is achieved by assigning an identifier to a digital object, i.e. as d.o.i works for articles.

CHAIN-REDS TOOLS

In this Section, we will present a brief description of the CHAIN-REDS tools that are available at its website and are of interest to Data Infrastructure. They actually are the backbone that is being used by the DART challenge to achieve data trust building and will be the basis for several of the use cases coming from the regions.

THE CHAIN-REDS KNOWLEDGE BASE

The CHAIN-REDS Knowledge Base is one of the largest existing e-Infrastructure-related digital information systems. It currently contains information, gathered both from dedicated surveys and other web and documental sources, for largely more than half of the countries in the world.

In principle, the KB was implemented considering e-Infrastructure as an environment where research resources (hardware, software and content) can be readily shared and accessed where necessary to promote better and more effective research. Then, such environments integrate hard-, soft- and middleware components, networks, data repositories, and all sorts of support enabling virtual research collaborations with a final goal: to allow scientists across the world to do better (and faster) research using DCIs, independently of where they are and of the paradigm(s) adopted to achieve their goals. Thus, to better fulfill this long term milestone, the use of standards is more than an asset and, at the same time, a step forward to achieve sustainability.

The first release of the CHAIN KB presented integrated dynamically updated information about DCIs. Information about Regional and National Research and Education Networks, National Grid Initiatives, Certification Authorities, Identity Federation Providers, Regional Operation Centers, Grid sites and Applications (and already running on a Science Gateway) was available in both Country and Table views. This on-line service was a clear step forward in harmonizing the different regional infrastructure information, new capabilities should be incorporated. Once standards were identified to easily gather and access both OADRs and DRs, a demonstrator was built with them to visualize and access the repositories by means of both geo- and tab-views (as it was previously made for DCIs). Such a demonstrator was implemented with the advances carried out within the Knowledge Linking and sharing in research dOmainS (KLIOS) project. KLIOS is based on the interconnection and the integration of scientific resources through a grid of meta-data network and provide the following services: metadata harvesting; semantic enrichment; and, linked data semantic search.

Basically, the new KB capability is composed of a multi-layer structure where two harvesters running on either Grid or Cloud search for OAI-PMH endpoints from OADRs and DRs. Above them, a semantic web-enrichment layer is used to act as a previous step before the linked-data search engine, which is on the top. The process of the metadata harvesters is as follows:

- Get the address of each repository publishing an OAI-PMH standard endpoint
- Retrieve, using the OAI-PMH repository address, the related Dublin Core encoded metadata in XML format
- Get the records from the XML files and, using the Apache Jena API, transform the metadata in RDF format



Figure 1. A snapshot of the CHAIN-REDS Knowledge Base - OADR Site view (red markers refer to repositories found by automatic harvesting and yellow ones to those directly integrated by means of CHAIN-REDS activities).

Basically, the new KB capability is composed of a multi-layer structure where two harvesters running on either Grid or Cloud search for OAI-PMH endpoints from OADR and DRs. Above them, a semantic web-enrichment layer is used to act as a previous step before the linked-data search engine, which is on the top. The process of the metadata harvesters is as follows:

- Get the address of each repository publishing an OAI-PMH standard endpoint
- Retrieve, using the OAI-PMH repository address, the related Dublin Core encoded metadata in XML format
- Get the records from the XML files and, using the Apache Jena API, transform the metadata in RDF format

Save the RDF files into a Virtuoso triple store according to an OWL-compliant ontology built using Protégé.

As it has been aforementioned, as a legacy from CHAIN, it has 86 entries in its DCIs Table view, which reports on the country where the DCI is settled, the regional network, the National Research Education Network (NREN) and the National Grid Initiative (NGI) it belongs to, the Certification Authority (CA) and the Identity Federation (IdF) it relies on for accessing it, the Regional Operation Centre (ROC) it is connected to and the sites it counts on. As in the previous case, the same information can be showed searching on a world map.

THE CHAIN-REDS SEMANTIC SEARCH ENGINE

Even when the KB is searchable in one of the four topics it contains (country, name, domain, organization), the CHAIN-REDS consortium has decided to semantically enrich the OADR and DRs gathered in the KB and build a search engine on the related linked data. The CHAIN-REDS Semantic Search Engine has been the result of such an effort.

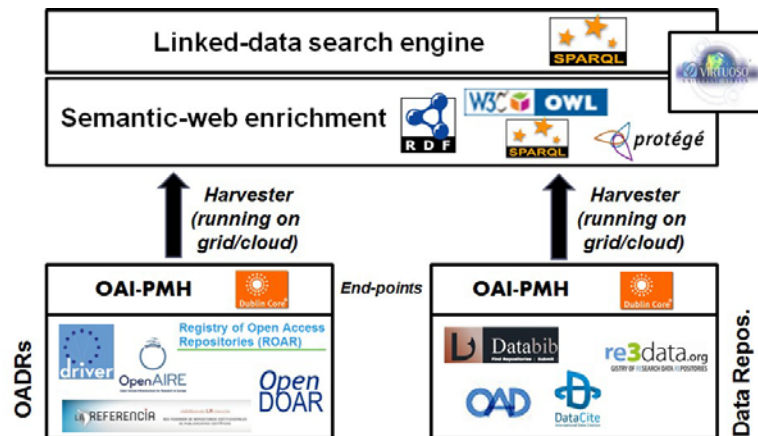


Figure 2. Architecture of the Semantic Search Engine

The multi-layered architecture of this engine is sketched in Figure 2 where both the official and *de facto* Semantic Web standards and technologies adopted are described by small logos.

Using it, visitors can either enter a keyword and submit a SPARQL query to the Virtuoso triple store or select a language and get, on the left side of the page, the list of subjects available in that language with the indication, between parentheses, of the number of records available for that particular subject (see Figure 3).

The results of a given query are listed in a summary view directly displayed on the webpage. For each record found, the title, the author(s) and a short description of the corresponding resource are provided. Clicking on the “More Info” link, visitors can access the detailed view of the resource. In the “Dataset information” panel users get the link to the open access document and, if existing, to the corresponding dataset. Clicking on the “Graphs” tab, which appears at the top of the summary view, users can select one or more of the resources found and get a graphic view of the semantic connections among Authors, Subjects and Publishers, as shown in Figure 4. In this way, if new links appear, connecting different resources (as shown in the lower left corner of the figure), users can infer new relations among resources, thus discovering new knowledge.

The technological description of how this process can be made follows. The first pillar is the harvester procedure. Then, each Resource Description Framework (RDF) file retrieved and saved in a Virtuoso-enabled triple store is mapped onto a Virtuoso Graph that contains the ontology expressly developed for the search engine. The ontology, built using Dublin Core and FOAF standards, consists of:

- Classes that describe the general concepts of the domain: Resource, Author, Organization, Repository and Dataset (where Resource is a given open access document);

- Object properties that describe the relationships among the ontology classes; the ontology developed for the service described in this paper has several specific properties such as `hasAuthor` (i.e., the relation between Resources and Authors) and `hasDataSet` (i.e., the relation between Resources and Datasets)
- Data properties (or attributes) that contain the characteristics or classes parameters.

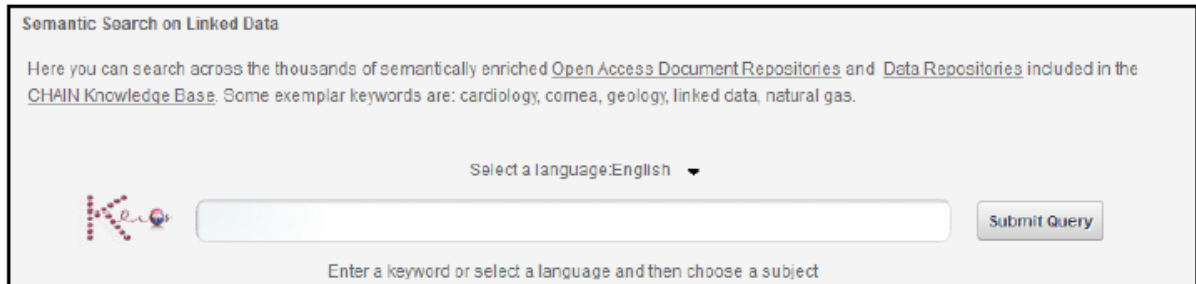


Figure 3. Schema of the ontology used for the Semantic Search Engine

The highest-level, component is the Search Engine itself. Using it, visitors can either enter a keyword and submit a SPARQL query to the Virtuoso triple store or select a language and get, on the left side of the page, the list of subjects available in that language with the indication, between parentheses, of the number of records available for that particular subject. The results of a given query are listed in a summary view directly displayed on the webpage. For each record found, the title, the author(s) and a short description of the corresponding resource are provided. Clicking on the “More Info” link, visitors can access the detailed view of the resource.

In the “Dataset information” panel users get the link to the open access document and, if existing, to the corresponding dataset. Clicking on the “Graphs” tab, which appears at the top of the summary view, users can select one or more of the resources found and get a graphic view of the semantic connections among Authors, Subjects and Publishers. In this way, if new links appear, connecting different resources, users can infer new relations among resources, thus discovering new knowledge.

A programmable use of the CHAIN-REDS SSE is also possible due to the development of a RESTful API that has been created on purpose; now, it is possible to get and/or re-use the many millions of open access resources contained in the CHAIN-REDS KB and stored in a Virtuoso RDF-compliant database by calling the Semantic Search Engine from a common website or even mobile application.

Now it is possible to perform either single or parallel semantic searching¹³ [26]. By passing the mouse over the "Semantic Search" link of the CHAIN-REDS webpage, any user can see a sub-menu with several items; the first two are:

- Single: the usual semantic search service described above; and,
- Parallel: the new parallel semantic search service that allow users to search in parallel (i.e., at the same time) across the millions of resources contained in the CHAIN-REDS Knowledge Base and in the ENGAGE Platform¹⁴. Parallel semantic search engines have been made available also in the SGs of some (collaborating) projects, enhancing and extending in this way the solutions proposed by CHAIN-REDS. This parallel semantic search can be found at:
- agINFRA, here the user can search in parallel across the millions of resources contained in the CHAIN-REDS Knowledge Base and in the OpenAgris¹⁵ repository; and,
- DCH-RP, here the user can search in parallel across the tens of millions of resources contained in the CHAIN-REDS Knowledge Base and in the Europeana¹⁶, Cultura Italia¹⁷ and Isidore¹⁸ repositories.

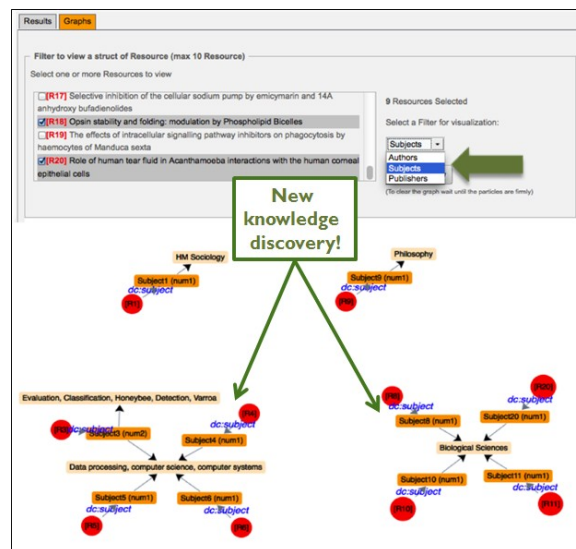


Figure 4. Graphic connections among records found by the Semantic Search Engine

Another two extensions have been the programmable use of the Semantic Search Engine by a RESTful API and the implementation of the engine on a mobile app for both Apple store and Google Play. Any user can access them also by passing the mouse over the "Semantic Search" link of the CHAIN-REDS webpage.

THE DART CONCEPT

The vision of the emerging Data Centric Science is that a researcher, of a given scientific domain, could find publications and being automatically redirected to the data used to produce those papers and to the applications used to produce those data. Or, alternatively and simply, access raw data of interest to be later used as input in applications.

Researchers can then access that data and the corresponding application in order to reproduce and extend the results of a given study. The new data (and the new paper if any) are stored on the Data Infrastructure and can be easily found by the people belonging to the same domain making possible to start the cycle again (see the diagram below in Fig 5). The requirements that are needed and not directly managed by CHAIN-REDS are related to intellectual properties issues and unique identifiers (PID) referring to papers, data and applications. Nevertheless, CHAIN-REDS is supporting the assignment of PID to digital objects by means of the service provided by the partner GRNET¹⁹.

CHAIN-REDS has started working on this DART challenge for providing proof-of-principle use-cases for data sharing across continents. The first prototype has been successfully tested. It counts on a couple of datasets stored in ZENODO²⁰ and DataCite²¹ and applications that could make use of them. Thus, datasets related to molecular cross sections hosted by the Max-Planck Institute²² can be downloaded from ZENODO and data on genes to be compared across species hosted by Ensembl²³ can be downloaded from DataCite. The former can be used to obtain molar absorption coefficients by using an application devoted by CIEMAT already included in a Science Gateway portlet called Molon²⁴ and the latter can be taken as input for several applications, like for example jModelTest²⁵, to obtain models of nucleotide substitution.

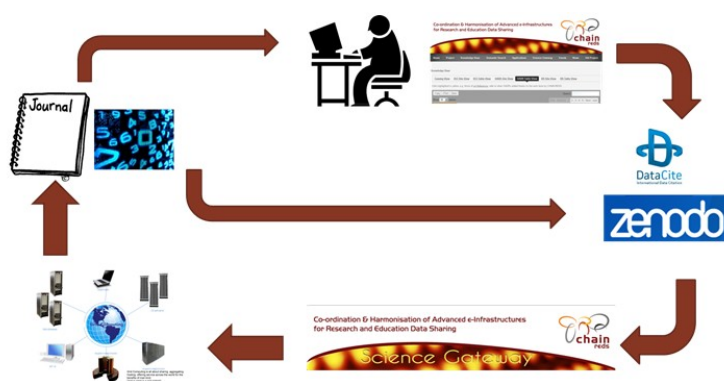


Figure 5. The DART challenge.

This proof-of-concept test is going to be shown and proposed to the CHIAN-REDS collaborative communities for their own use with their own datasets and applications. These communities come from the Agriculture, e-Government, Earth Science, Cultural Heritage and Astroparticle domains. This concept is also being promoted in the CHAIN-REDS targeted regions in order to find success stories; Latin America the LAGO (Latin American Giant Observatory) Collaboration has been identified, and will be described below. At the same time, DART has been proposed to the Workflows Working Group of the EUDAT initiative²⁶ and now CHAN-REDS representatives are part of it.

LAGO AS A DART EXAMPLE

The Latin America Giant Observatory (LAGO) (first known as *Large Aperture Gamma Ray Observatory*²⁷) project is a recent collaboration that comes from the association of more than 80 Latin American astroparticle researchers, keeping a close collaboration with researchers at IN2P3 in France and INFN in Italy. This collaboration was motivated by the experience of the Pierre Auger Observatory, and the idea to install Water Cherenkov Detector (WCDs) in 9 Latin American countries: Argentina, Bolivia, Brazil, Colombia, Ecuador, Guatemala, Mexico, Peru and Venezuela. It started in 2005 and it was originally designed to survey the high-energy component of GRBs. Today it is a network of ground-based WCDs, located at different altitudes from Mexico through Antarctica (see Figure 6 below), devoted to study space climatology effects and GRBs signals on ground-based detectors by measuring the variations of the flux of secondary particles at ground level. Long-term modulation and transient events can also be characterized by using the LAGO detection network, as it spans over a big area with different sites at different latitudes, longitudes and geomagnetic rigidity cut-offs. Presently LAGO collaboration has 10 up/running WCD and it is planned double them in the next two years, with five more new detectors installed in 2014 and other five in 2015. Typically, each detector generates 150 GB of data per month and the entire collaboration generates 1.5 TB/month. This experimental data is preserved locally and shared through a data repository based on DSpace²⁸.

Additionally, at each site the particle flux over the ground detector has to be simulated and correlated to the signals emerging from the detectors. The particle flux simulation is carried out using CORSIKA²⁹ (COsmic Ray Simulations for KAscade) a software for detailed simulation of extensive air showers initiated by high energy cosmic ray particles. CORSIKA is extensively used by the astroparticle community. Particularly, CORSIKA is used by the Auger Observatory³⁰ collaboration and more recently the HAWC³¹ project. Typically CORSIKA simulations generate 10GB/site. These synthetic data is also preserved in the data repository. The collaboration also uses GEANT4³² to evaluate the response of the instrumentation of the WCD to the crossing of particles through the water volume of the detector.

LAGO is an experiment that could handle, with reasonable scale, a distributed community, collaborating across Latin America, building a network of data repositories through the continent, using computational intensive applications and developing an outreach program to promote Data Science as a Citizen Science initiative.

Inspired by the debutante, the Open Data Movement, LAGO has been developing four main initiatives to deploy an open network of curated data repositories, namely:

1. To preserve, curate and share the data registered by WCDs array, now through data repository³³ (DR) and in the near future across a network of DR
2. To generate a toolkit of scripts and algorithms to detect the proper operation of the detectors. This toolkit will part of the next generation of firmware and will allows the system to reconfigure some of it parameters in order to minimize some of diagnosed malfunctioning.
3. To offer a computational infrastructure that allows the collaborating members to ubiquitously analyse the curated data efficiently
4. To open part of the data and use it to outreach Data Science to university students.

These initiatives aim to pave the way to openly share the data recorded by LAGO collaboration with any other domain disciplines and to Citizen Science initiatives.

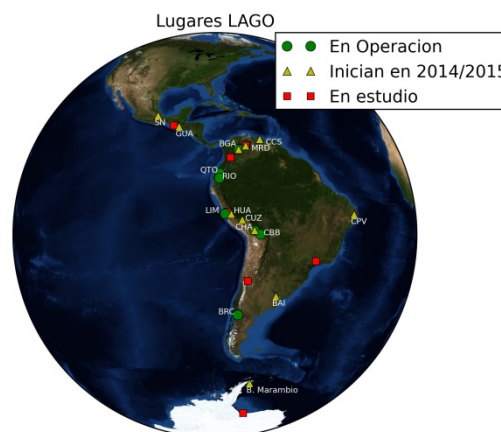


Figure 6. LAGO Sites: Green in operation, starting operation on 2014/2015 and red under study/consideration.

LAGO has developed a prototype of data repository, LAGOData³⁴ as part of a more ambitious project, LAGOVirtual³⁵ oriented to develop a working environment to have access and analyse data recorded in all LAGO Sites. In LAGO repository data is classified into three types: instrument calibration data, WCD data sets and simulated data sets. In the near future we want the members of the collaboration to use this repository also to preserve papers, thesis, Labs Notes and Technical Reports related to the project. The idea is to link documents to all the data sets analysed in it.

Each data file is tagged by a metadata set specifically adapted to LAGO. The existence and implementation of a scientific metadata standard model will allow an uniform access to data for all the LAGO collaboration members, the interoperability between scientific information systems and also will contribute to the data preservation and its usability in time. The metadata model the collaboration uses for LAGOData is an adaptation of the model raised by the Council for the Central Laboratory of the Research Councils³⁶.

LAGO repository exposes data/metadata through the Open Archives Initiatives Protocol for Metadata Harvesting³⁷. This protocol is used by external systems to collect the data and metadata and create aggregated value services like meta- searchers.

We expect to release a second version of LAGOData by the end of 2014 which will have an implementation of the SWORD (for Simple Web-service Offering Repository Deposit) protocol which became a way to address the need for a standardised deposit interface to digital repositories³⁸.

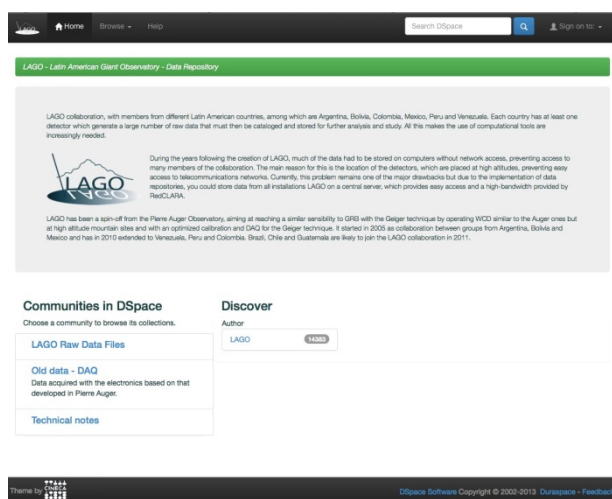


Fig. 2. Screen shot of LAGOData repository <http://halley.uis.edu.co/LAGOVirtual/>



An important Data Science outreach activity is carried out by using LAGO Data Repository. LAGO Colombia is developing an intense outreach program on Data Science, teaching statistical data analysis to university students. This initiative is supported by ColCiencias (Colombian Science and Technology funding agency)

Presently, LAGO collaboration has COSIKA and GEANT4 installed at the Universidad de Santander Supercomputing Center³⁹ and in a small dedicated cluster of a six Workstation (Quad Core Intel Xeon E5520, 4 GB RAM/node and 2 TB DD/node) with NVIDIA Quadro NVS 420 which controls a visualization wall of 16 monitors of 24 inches, capable to generate a resolution of 32 MP. It is foreseen to have access to the CORSIKA grid version which is available in other Astroparticle Observatories and which is supported by the ROC-LA.

CONCLUSIONS

CHAIN-REDS has as a major goal to propose a model for accessing and managing data. To achieve such an objective, several tools based on standards have been implemented and access to these services through identity providers has been relied on. Now, any user can perform a whole cycle of searching data and documents, retrieve the raw data, use them as an input of a service, and obtain final results susceptible of being stored under the same standards format.

For the specific case of Latin America, CHAIN-REDS is closing collaborating with RedCLARA. Thus, it is supporting the SCALAC service and integrating major repositories as La Referencia into the project Knowledge Base. Also, success stories are being aware of the CHAIN-REDS developments and will benefit from them.

In principle, scientific cases have been identified, but all the services promoted by CHAIN-REDS can be adopted by the Academia and the NRENs. In this sense, CHAIN-REDS is much interested in collaborating with this kind of institutions to support them in the use of its services.

ACKNOWLEDGEMENTS

This work has been partially funded by the European Commission Seventh Framework Programme project “Co-ordination & Harmonisation of Advanced e-Infrastructures for Research and Education Data Sharing” (CHAIN-REDS, Grant agreement 306819).

REFERENCES

1. See a classical discussion of this problem of how small labs and institutes are facing the data deluge problems in C Borgman, J Wallis, and N Enyedy.(2007) **Little science confronts the data deluge: habitat ecology, embedded sensor networks, and digital libraries** *Int J Digit Libr*, **7**, 17-30.
2. Arzberger, P., Schroeder, P., Beaulieu, A., et al. (2004) **Science and government: an international framework to promote access to data**. *Science*, **303**:1777–1778; Simberloff, D., Barish, B. C., Droegemeier, K. K., et al. (2005) **Long-lived digital data collections: enabling research and education in the 21st century**. Technical Report NSB-05-40, National Science Foundation, Washington DC, USA; and Lyon, L. (2007) **Dealing with data: roles, rights, responsibilities and relationships**. Consultancy Report, UKOLN, University of Bath, UK.
3. See **Report of The European Commission Public Consultation on Open Research Data** (July 2013) http://ec.europa.eu/research/science-society/document_library/pdf_06/report_2013-07-open_research_data_consultation.pdf and **Digital Research Data Sharing and Management** Committee on Strategy and Budget National Science Board NSF Technical Reports (2011) <http://www.nsf.gov/nsb/publications/2011/nsb1124.pdf>
4. CHAIN-REDS interoperability demo, <http://science-gateway.chain-project.eu/demo-status> and also CHAIN-REDS D4.2, available at <http://www.chain-project.eu/deliverables>
5. CHAIN-REDS Knowledge Base, <http://www.chain-project.eu/knowledge-base>
6. OAI-PMH, <http://www.openarchives.org/pmh/>
7. Dublin Core, <http://dublincore.org>
8. SPARQL, <http://www.w3.org/2001/sw/wiki/SPARQL>
9. XML, <http://www.w3.org/XML/>
10. PID, <http://www.pidconsortium.eu/>
11. KLIOS, <http://klios.ct.infn.it>
12. The programmable use of the Semantic Search Engine, <http://www.chain-project.eu/semantic-search-api>
13. The parallel search in the CHAIN-REDS KB, <http://www.chain-project.eu/parallel-semantic-search>



14. ENGAGE, <http://www.engagedata.eu/>
15. The OpenAgris repository, <http://aims.fao.org/openagris>
16. The Europeana repository, <http://www.europeana.eu/>
17. The Cultura Italia repository, <http://www.culturaitalia.it/>
18. The Isidore repository, <http://www.rechercheisidore.fr/>
19. GRNET PID Service, <http://epic.grnet.gr/>
20. TheZENODO repository, <http://zenodo.org/>
21. The DataCite initiative, <http://www.datacite.org/>
22. The MPI-Mainz UV/VIS Spectral Atlas, http://satellite.mpic.de/spectral_atlas
23. Ensembl Gene Trees and Homologues,
<http://www.ensembl.org/info/website/tutorials/compara.html>
24. Molon portlet, <http://science-gateway.chain-project.eu/molon>
25. M. Loureiro et al. Grid selection of models of nucleotide substitution. Studies in Health Technology and Informatics 159, 244-248 (2010) and also in jModelTest2 portlet,
<http://science-gateway.chain-project.eu/jmt>
26. EUDAT, <http://www.eudat.eu/workflows>
27. Fullauthor/institution list at <http://www.lagoproject.org>
28. An open source software that enables open sharing of many types of content, generally used for institutional repositories <http://www.dspace.org>
29. <https://web.ikp.kit.edu/corsika/>
30. <http://visitantes.auger.org.ar/index.php/el-observatorio/colaboracion-internacional.html>



32. <http://www.inaoep.mx/~hawc/>
33. The standard a toolkit to simulates the passage of particles through matter. See <http://geant4.cern.ch>
34. <http://halley.uis.edu.co/LAGOVirtual/>
35. L.A. Torres, L.A. Núñez, R. Torrén, and E.H. Barrios. Implementación de un repositorio de datos científicos usando dspace. *E-Colabora*, 1(2):101–117,
36. 2011
37. R. Camacho, R. Chacón, G. Díaz, C. Guada, V. Hamar, H. Hoeger, A. Melfo, L. A. Núñez, Y. Pérez, C. Quintero, M. Rosales, and R. Torrén. Lagovirtual. a
38. collaborative environment for the large aperture grb observatory. In R. Mayo, H. Hoeger, L. Ciuffo, R. Barbera, I. Dutra, P. Gavillet, and B. Marechal, editors,
39. *Proceedings of the Second EELA2 Conferencem Choroní Venezuela*, Madrid
40. España, 2009. EELA2, CIEMAT.
41. Shoaib Sufi and Brian Mathews. Cclrc scientific metadata model: Version 2.
42. *Final report. Council for the Central Laboratory of the Research Councils. Report No: DL-TR-2004-001*, 2004.
43. <http://www.openarchives.org/pmh/> also Herbert Van de Sompel, Michael L
44. Nelson, Carl Lagoze, and Simeon Warner. Resource harvesting within the oai- pmh framework. *D-lib magazine*, 10(12):1082–9873, 2004
45. Stuart Lewis, Pablo de Castro, and Richard Jones. SWORD: Facilitating
46. Deposit Scenarios. *D-Lib Magazine*, 18(1-2), January 2012.
47. <http://sc3.uis.edu.co>



Preservação digital

Políticas e planos de preservação, ferramentas para preservação estratégias e implementação,
curadoria de conteúdos, futuro e tendências da preservação



DESAFÍOS DE LA PRESERVACIÓN DIGITAL DE LOS ARCHIVOS SONOROS

Dra. Perla Olivia Rodríguez Reséndiz

Investigadora

Instituto de Investigaciones Bibliotecológicas y de la Información

Universidad Nacional Autónoma de México

Dirección: Circuito Interior s/n, Torre II de Humanidades, piso 12, Ciudad Universitaria, C.P. 04510,
Delegación Coyoacán, México, D.F.

Correo electrónico: perla@iibi.unam.mx

RESUMEN

La música, los programas de radio, las voces de poetas, escritores, científicos, políticos; los testimonios de pueblos originarios, el paisaje sonoro, el arte sonoro, entre otros tipos de grabaciones que se resguardan en archivos sonoros, son evidencias del devenir histórico, fuentes de información, medios para la adquisición de conocimiento, vehículos de educación y bienes culturales que constituyen parte del patrimonio sonoro de la humanidad. Se estima que existen más de cien millones de horas documentos sonoros que han sido grabados en diversos soportes sonoros y que están en riesgo de perderse en los próximos años. De hecho, investigaciones recientes dan cuenta de que cada década se perderá el 30% del patrimonio sonoro en Europa y el 50% en América Latina. En la era de la información digital, caracterizada por la generación diaria de miles de documentos digitales, coexiste la inevitable pérdida de documentos sonoros que durante más de 150 años han sido grabados en soportes analógicos. Ante tal escenario, la única forma de garantizar la preservación de los documentos sonoros es digitalizarlos y conservarlos en plataformas digitales. En consecuencia, en esta ponencia se analizan los desafíos que presenta la preservación digital para garantizar la conservación y acceso de los contenidos de archivos sonoros. Estos desafíos son:

- Conocimiento de la cantidad, contenido y condiciones en que se encuentran los archivos sonoros
- La cooperación y colaboración sin fronteras ante deterioro de los archivos sonoros
- La incorporación de documentos sonoros de origen digital en el archivo
- La sustentabilidad del archivo digital sonoro

Palabras clave: Preservación digital, preservación digital sonora, preservación sonora, archivos sonoros.



ABSTRACT

The music, the radio programs, the voices of poets, writers, scientists, politicians; the testimonies of indigenous peoples, the soundscape, sound art, among others types of sound recordings are evidence of the historical development, information source, mediums for acquiring knowledge, education and cultural vehicles property constituting part of the sound heritage humanity. There are over one hundred million sound recordings that have been recorded in various sound supports in risk of being lost in the coming years. In fact, recent research realized established that every decade will miss 30% of the sound heritage in Europe and 50% in Latin America. In the age of digital information, characterized by the daily generation of thousands of digital documents, coexists the loss of sound archives that for over 150 years have been recorded on analog media. The only way to guarantee the preservation of the sound archives is digitized documents and keeps them in digital platforms. In fact, in this paper are analyzed the challenges of digital preservation to ensure the sound conservation and access to sound archives contents. These challenges are:

- Knowledge of the quantity, content and conditions of the sound archives
- Cooperation and collaboration without borders front to deterioration of the sound archives
- The incorporation of sound document of digital source in the archive
- Sustainability of the digital sound archive

Key words: Digital preservation, digital sound preservation, sound preservation, sound archives.

INTRODUCCIÓN

La preservación de documentos sonoros grabados en soportes analógicos establece que el tiempo de vida de todos los soportes de audio está determinado por su estabilidad física y química, es decir, por contar con adecuadas condiciones de conservación, manipulación y por minimizar el uso del documento en el acceso público. Por ello, durante muchos años el trabajo en los archivos sonoros se centró en procurar condiciones de conservación y en crear copias para el acceso, como una medida de preservación de los originales en el archivo (Rodríguez, 2012).

Con el advenimiento de la tecnología digital y su incorporación en los archivos sonoros esta situación cambió. Ya no fue necesario hacer copias en soportes analógicos para conservar los documentos sonoros. La digitalización de contenidos de diversos soportes analógicos significó un medio para conservar los materiales sonoros y ensanchó sus posibilidades de acceso; además, inauguró una nueva era de información, conocimiento y memoria (Hoog, 2005); fue el primer paso para la preservación digital de contenidos sonoros.



Schuller (2008) ha dicho que entre 1989 y 1990, el paradigma de conservación de soportes originales en los archivos sonoros cambió. Los responsables de los archivos sonoros se dieron cuenta que conservar los materiales originales era un esfuerzo vano, debido a la inestabilidad y fragilidad de los soportes analógicos y a la obsolescencia de los equipos de grabación y reproducción sonora. Por ello, la estrategia más importante fue conservar el contenido sonoro grabado en diversos soportes sonoros. La preservación se centró en los contenidos sonoros para que una vez transferidos, a través de la digitalización, a plataformas digitales, fueran conservados a través de copias digitales de mayor fidelidad (loss less copying) con el documento original. Conviene precisar que la preservación digital, como lo ha señalado Ray Edmondson (2010) no debe ser entendida como sinónimo de copias digitales. La preservación digital es parte de un concepto más complejo que garantiza la integridad y el acceso de los documentos sonoros. La preservación digital puede ser definida como el método sustentable para conservar, administrar, gestionar, dar acceso, difusión y reaprovechamiento de contenidos sonoros a través de plataformas digitales.

ANTECEDENTES HISTÓRICOS

En 1992 se llevó a cabo el primer proyecto de digitalización de un archivo sonoro, en la Arbeitsgemeinschaft der Öffentlich-rechtlichen Rundfunkanstalten der Bundesrepublik Deutschland/ Consorcio de instituciones públicas de radiodifusión de la República Federal de Alemania (ARD). A partir de entonces, el interés por la digitalización de colecciones sonoras llamó la atención e interés de los responsables de los archivos sonoros, motivó la reflexión y la polémica.

En un inicio, algunos archivistas y documentalistas se mostraron renuentes a llevar a cabo proyectos de digitalización debido a que consideraban que estaban ante un nuevo ámbito de conocimiento profesional para el que no estaban capacitados y que ponía en riesgo su estabilidad laboral; además, la inestabilidad del formato digital, suponía un riesgo para la preservación a largo plazo. Por ello, la formulación de estándares y recomendaciones para la digitalización de archivos sonoros constituyó una significativa aportación para garantizar que la transferencia de contenidos sonoros a plataformas digitales fuera realizada bajo las condiciones tecnológicas adecuadas y en los formatos que garantizaran la integridad de la información. Algunas de las más importantes recomendaciones y estándares a considerar establecieron que la digitalización debe realizarse sin compresión, los formatos recomendados son WAV o BWF; el muestreo debe de ser de 48 Khz y 24 bits cuando menos de cuantificación; es necesaria la recuperación óptima de la señal analógica del documento sonoro con los equipos adecuados de grabación y reproducción sonora; la digitalización debe llevarse a cabo sin alterar el documento original y se debe establecer el vínculo de la media y la metadata a través del ID (IASA, 2003 y la ARSC, 2011).



]

Por otra parte, la digitalización de colecciones sonoras, también provocó una fascinación que motivó el uso de CD (discos compactos) y DVD como soportes de almacenamiento para tener una copia de los documentos digitalizados. Esta práctica, se alentó por el bajo costo de los nuevos soportes digitales; sin embargo, en la elaboración de estas copias digitales se omitieron, en la mayoría de los casos, por desconocimiento, las recomendaciones y estándares que garantizan la integridad de la información.

De forma simultánea al embeleso por tener una copia del archivo en soportes digitales, se llevaron a cabo reuniones e investigaciones de expertos en relación con el uso de discos compactos y DVD como soportes de almacenamiento (Rodríguez, 2012). El investigador australiano Kevin Bradley (2007) fue uno de los primeros expertos en reconocer que los CD y los DVD son los soportes más inestables y frágiles en la historia de la grabación sonora, por lo que no se recomiendan como medios de almacenamiento. Bradley (2007) estableció que el decaimiento y la obsolescencia del soporte, demandarán que sea migrado a otro soporte en el futuro, por lo cual, teniendo en consideración estas limitaciones, se pueden hacer copias digitales por un periodo limitado de tiempo, y de acuerdo a lineamientos y estándares. Ante la imposibilidad de guardar los contenidos digitalizados en CD y DVD, fue necesario pensar en un soporte que garantizara la conservación a largo plazo.

En consecuencia, el almacenamiento masivo digital se erigió como un ámbito de interés que de forma progresiva fue ganando presencia. Al respecto, Schuller (2008) señaló que los Sistemas de Gestión y Almacenamiento Masivo Digital (SGAMD) fueron discutidos por primera vez, a inicios de la década de los noventa, por los archivistas sonoros y comenzaron a realizarse proyectos piloto para explorar el potencial de estos sistemas para el almacenamiento sonoro. A mediados de los años 90, algunos archivos sonoros comenzaron a adquirir (SGAMD), cuyo costo de adquisición era muy alto.

En ese momento, la preservación del patrimonio sonoro en plataformas digitales se tornó un aspecto crítico, que motivó la discusión internacional en relación con la necesidad de encontrar opciones para que los archivos pequeños y con pocos recursos económicos, también pudieran contar con un SGAMD. Esta discusión formó parte de la polémica en relación con la brecha digital que divide a los países ricos de los pobres. La inequitativa situación de acceso a la tecnología fue el tema abordado en la conferencia *Bridging the digital divide by providing support to content professionals in the least Developer countries* organizada por la UNESCO en 2003.

En respuesta a la preocupación motivada por el riesgo de perder miles de horas que forman el patrimonio sonoro de la humanidad, se impulsaron iniciativas, proyectos y publicaciones con el objetivo de buscar alternativas para la preservación del patrimonio sonoro en plataformas digitales.



En Europa, se creó el proyecto Presto Space con fondos de las principales radiodifusoras y televisoras públicas europeas, como son la BBC de Londres, el Instituto Nacional del Audiovisual de Francia, la Radio Televisione Italiana, entre otras; para buscar soluciones técnicas y sistemas integrales de preservación digital de colecciones sonoras y audiovisuales. Además, el investigador australiano, Kevin Bradley propuso el diseño de un sistema simple y sustentable de preservación digital a partir de uso de software de código abierto.

Por lo tanto, desde hace más de dos décadas, la preservación digital sonora constituye un nuevo ámbito que se ha incorporado en el trabajo cotidiano de los repositorios sonoros. Aun cuando ya se cuenta con estándares y recomendaciones para la adecuada digitalización de colecciones sonoras y se han desarrollado sistemas de gestión y almacenamiento masivo digital para la administración, conservación y acceso de colecciones sonoras, la preservación digital tiene ante sí cuatro desafíos que deberá enfrentar en un futuro inmediato.

Primer desafío: Conocimiento de la cantidad, contenido y condiciones en que se encuentran los archivos sonoros.

Fue hasta 1980, que los archivos sonoros y audiovisuales comenzaron a ser parte de la discusión y del interés público. En esta década, la UNESCO publicó la *Recomendación sobre la salvaguarda y la conservación de las imágenes en movimiento*, aprobada el 27 de octubre de 1980, y consideró que los documentos audiovisuales, en los cuales se incluyen los sonoros, forman parte del patrimonio intangible, son un bien cultural, son parte de la memoria y legado de la humanidad y conservarlos es una tarea de alta prioridad (UNESCO, 1980). A partir de entonces, las grabaciones sonoras en diversos soportes fueron apreciados por su valor patrimonial y por ser un legado para la humanidad. Sin embargo, el ámbito de los archivos sonoros era prácticamente desconocido. Se carecía de estudios que dieran cuenta de las condiciones de los documentos sonoros.

Fue hasta la década de los años 90, que se llevaron a cabo las primeras encuestas destinadas a conocer las condiciones de conservación de diferentes tipos de soportes sonoros analógicos como son: cilindros, discos de acetato, vinil, cassetes, cartuchos, cintas de carrete abierto, discos compactos, entre otros. En 1995, el Comité Técnico de la Asociación Internacional de Archivos Sonoros y Audiovisuales (IASA) desarrolló la primera encuesta para conocer el estado de conservación de 32 tipos de soportes sonoros, en 41 países de Europa y de América (Boston, 2003), de este último continente, sólo participaron los Estados Unidos, Canadá y Barbados. Los resultados de esta encuesta dieron a conocer que el 56% de los soportes conservados de 80 archivos sonoros estaban en condiciones de deterioro. Este resultado se constató con una segunda encuesta que realizó George Boston en 2003, además, el estudio se subrayó el riesgo de pérdida de documentos sonoros.



Luego de estos dos estudios que alertaron en relación con el riesgo de pérdida de archivos sonoros; en 2004, se señaló que la memoria audiovisual de la humanidad está formada por 200 millones de horas de material, de los cuales el 50% son documentos sonoros (Wright, 2011). Es decir, que 100 millones de horas de documentos sonoros forman el patrimonio sonoro de la humanidad. Esta cifra es, probablemente, el primer dato cuantitativo de la cantidad de documentos sonoros que se han grabado desde hace más de un siglo y medio. Además se señaló que el 80% de la memoria audiovisual, estaba en peligro de perderse, principalmente en los países en desarrollo.

Por otra parte, la encuesta TAPE realizada en 2007 por Klijn y de Lusenet (2008) identificó 9.4 millones de horas de audio que forman parte de las colecciones sonoras de 374 instituciones en 34 países europeos. Un dato relevante derivado de este estudio fue dar a conocer la existencia de 3.7 millones de horas que pertenecen a radiodifusoras. Asimismo, se publicó que aproximadamente la mitad de las colecciones sonoras son pequeñas, sus colecciones no ascienden a más de 500 horas con documentos valiosos. La encuesta destacó que la radio es una de las instituciones que detentan el mayor volumen de documentos sonoro. Además, en este estudio, los autores enfatizaron que los soportes sonoros conservados en las instituciones participantes en la encuesta no se encuentran en adecuadas condiciones de conservación y acceso. Por ello, se consideró como inevitable y urgente la transferencia de contenidos a soportes digitales, como la única forma de conservar y dar acceso a largo plazo los documentos sonoros (Klijn y de Lusenet, 2008).

Sam Brylawski y Rob Bamberger (2010) han señalado en un vasto y puntual estudio del patrimonio sonoro en los Estados Unidos, que sólo en ese país existen 46 millones de horas de documentos sonoros que pertenecen a 17 mil instituciones entre las que se incluyen bibliotecas, archivos, museos, asociaciones históricas, repositorios arqueológicos y colecciones de investigación. Se desconocen las condiciones en que se encuentra el 44% de las colecciones sonoras de estas colecciones. Asimismo, en el estudio también se reconoció que muchas grabaciones originales de la radio se han perdido, sobre todo de los primeros años del desarrollo de este medio.

El estudio establece que la digitalización se está realizando a través de diversas empresas de forma descentralizada. Llama la atención que en el estudio de Sam Brylawsky y Rob Bamberger (2010) se hace alusión a la presencia, cada vez mayor de documentos sonoros de origen digital. Al respecto, se señala que el riesgo de pérdida de un documento sonoro, no necesariamente está determinado por la edad de la grabación, sobre todo, porque en la actualidad miles de documentos digitales están en riesgo de desaparecer por carecer de sistemas adecuados de preservación.

Por otra parte, Presto PRIME realizó una encuesta con la participación de 400 universidades. Algunos resultados de la investigación dan cuenta de la existencia de 40 millones de horas en universidades de Estados Unidos y Europa, el 60% de estos documentos sonoros no está documentado, ni considerado en un plan de preservación. El 80% de estas colecciones audiovisuales tiene sólo una copia de sus materiales y están en riesgo de perderse (Wright, 2011).



Los datos derivados de los estudios en diversos países de Europa y Estados Unidos, proporcionan información cuantitativa de los documentos sonoros, en qué instituciones se resguardan y bajo qué condiciones de conservación. Los estudios han evidenciado el desconocimiento de las condiciones de conservación de los archivos sonoros y el riesgo de pérdida de los mismos, por la fragilidad de los soportes en que fueron grabados, la obsolescencia de tecnología para la grabación y reproducción sonora y la carencia de condiciones de conservación, entre otras variables. Es necesario destacar que las encuestas antes citadas han sido la base para sustentar, diseñar y justificar el desarrollo de proyectos de digitalización de colecciones sonoras y además han definido las estrategias a seguir para la preservación digital de colecciones sonoras.

Sin embargo, la situación de los archivos sonoros no es la misma en todo el mundo. En América Latina, se desconoce la cantidad y situación del patrimonio sonoro. Sólo se tiene evidencia que en 2003, se creó el Comité de Censo y Valoración de Registros Sonoros y Audiovisuales Etnográficos de los Países Andinos, integrado con representantes de Colombia, Ecuador, Bolivia, Perú y Venezuela, quienes diseñaron y aplicaron una encuesta para conocer el estado de los archivos etnográficos en esa región latinoamericana (IFEA, 2003). De acuerdo con los resultados de la encuesta los archivos sonoros de los países participantes no contaban con las condiciones adecuadas de conservación, carecían de inventarios, bases de datos y no tienen condiciones de acceso; además se identificó que las instituciones tienen muy poco presupuesto y adolecen de políticas que orienten la preservación del patrimonio sonoro. Se destacó el interés por reincorporar a las colecciones correspondientes los documentos que hayan sido expatriados; se evidenció la escasa consulta de los archivos sonoros y el nulo uso educativo del acervo sonoro (IFEA, 2003).

Observamos que los países en donde se han impulsado estudios e investigaciones para conocer la cantidad y condiciones de conservación de documentos sonoros han obtenido información que ha sido utilizada para diseñar, sustentar y poner en marcha proyectos y estrategias para la digitalización de documentos sonoros, así como para la generación de plataformas para la preservación digital sonora. En contrastante, los países donde se carece de estudios e investigación en relación con los archivos sonoros, se desconoce la situación de los archivos sonoros, y en consecuencia, se incrementa el riesgo de pérdida de este patrimonio. Sin el conocimiento de la cantidad, contenidos y condiciones de conservación de los documentos sonoros no es posible desarrollar propuestas de preservación digital sonora.

Por lo tanto, el primer desafío que se formula para la preservación digital de colecciones sonoras es la realización de estudios e investigaciones destinadas a conocer la cantidad, contenidos y condiciones en que se encuentran las colecciones sonoras que serán transferidas a plataformas digitales. En especial, para América Latina los estudios e investigaciones cuantitativas y cualitativas deben ser un instrumento para la definición de estrategias de preservación digital de los archivos sonoros.



Segundo desafio: La cooperación y colaboración sin fronteras ante deterioro de los archivos sonoros.

Se ha señalado que la digitalización de contenidos analógicos y su transferencia a plataformas digitales, es la única forma de asegurar la permanencia de miles de documentos sonoros que están en riesgo de perderse en las próximas décadas. Por ello, la digitalización puede ser entendida como una carrera contra el tiempo, contra el deterioro y pérdida de miles de documentos sonoros que forman parte del patrimonio sonoro y que han sido grabados en soportes analógicos.

La digitalización es una medida de preservación, a través de la cual se transfieren a plataformas digitales contenidos que han sido grabados en soportes analógicos. La digitalización es un proceso complejo, costoso y de largo plazo. El tiempo estimado para la digitalización de las colecciones depende de la cantidad de documentos, del tipo de soporte sonoro, las condiciones de conservación y catalogación de los soportes, la tecnología necesaria para la digitalización, el personal destinado a esta actividad y los recursos económicos otorgados a este fin.

Las instituciones que han iniciado proyectos de digitalización saben que el camino no es sencillo. Algunas experiencias pioneras de digitalización han tenido que enfrentar tropiezos y han tenido que aprender sobre la marcha. La experiencia acumulada es, sin lugar a dudas, una importante aportación para las instituciones que aún no han iniciado la transferencia de contenidos de documentos sonoros en soportes analógicos a plataformas digitales.

A más de dos décadas de haber iniciado el primer proyecto de digitalización de una colección radiofónica en Alemania, ya existen fechas para concluir la digitalización de colecciones sonoras. Estudios recientes señalan que la digitalización de las colecciones sonoras y audiovisuales en Europa terminará en el año 2073, tomando en consideración que cada año se digitalizan un promedio de 1.5% (0.28 millones de horas al año) (Addis, 2010). Sin embargo, el deterioro de los soportes sonoros es una constante que está presente en los planes de digitalización. Wrigth (2011) estimó, con base en cálculos del Image Permanence Institute relativos a la edad y condiciones de almacenamiento de las colecciones sonoras, que cada década el 30% de los documentos sonoros comenzará a ser ilegible como resultado de la obsolescencia y decadencia del material, es decir, se perderá más material del que se digitaliza. Lo que significa que en 2030, cuando el 33% del material se haya digitalizado el 67% será inutilizable. (Wrigth, 2011). Si se toma en consideración que un gran porcentaje de pequeños archivos carecen siquiera de las condiciones mínimas de conservación, el porcentaje de pérdida por década se puede incrementar hasta el 50% (Wrigth, 2011). Esta situación es crítica sobre todo para los archivos sonoros de emisoras de radio, coleccionistas particulares, colecciones de bibliotecas, de universidades y centros de investigación, entre otras, de América Latina, África y Asia.



Este escenario se agrava si se considera, como ha sido señalado, que salvo en Europa y Estados Unidos, no se han desarrollado estudios para conocer a detalle la cantidad y estado de conservación de los archivos sonoros. Al concluir la próxima década posiblemente los estudios se enfoquen en identificar los documentos que se han perdido y con ellos, una parte de nuestro patrimonio sonoro.

En consecuencia, otro de los desafíos más relevantes que enfrentan los archivos sonoros en el mundo, es diseñar estrategias de colaboración local, nacional y regional para la identificación, acopio y digitalización de documentos sonoros de valor social, histórico, político y cultural que estén en riesgo de perderse. Asimismo, es necesario subrayar la cooperación en torno a la investigación y la capacitación profesional sistemática que permita la actualización en conocimientos en relación con la preservación digital.

En este sentido, es necesario subrayar que ya existen iniciativas relevantes cuyo propósito es sumar su conocimiento y experiencia en la carrera contra el tiempo por preservar nuestra memoria sonora. Por ejemplo, el Instituto Max Planck es un centro de investigación que ofrece ayuda a los pequeños archivos, que no cuentan con la tecnología y los recursos económicos, para digitalizar grabaciones de lenguas de pueblos originarios en peligro de extinción.

Otro relevante modelo de cooperación es *Memoriav* fundada en 1995, como una red de instituciones y personas, para el intercambio de conocimiento técnico e información para la salvaguarda, valorización y divulgación del patrimonio sonoro y audiovisual de Suiza. Un aspecto de interés de esta propuesta es que incorpora colecciones de la Fonoteca Nacional de Suiza.

La pérdida de la memoria sonora es una realidad inminente ante la que los archivos, fonotecas, bibliotecas e instituciones que resguardan colecciones sonoras deberán analizar y determinar cuál es el mejor camino y la estrategia más adecuada no sólo para transferir contenidos analógicos a plataformas digitales, sino para crear plataformas de preservación digital sonora a largo plazo. Por ello, el segundo desafío es la creación de modelos de cooperación que beneficien a grandes y pequeños archivos sonoros.

Tercer desafío: La incorporación de documentos sonoros de origen digital en el archivo.

La digitalización de miles de documentos sonoros es una tarea monumental, una carrera contra el tiempo, para transferir a plataformas digitales contenidos sonoros grabados en soportes analógicos antes de se hayan deteriorado por completo. Por ello, gran parte de los recursos, infraestructura, personal y tiempo de un archivo sonoro está dedicado a esta tarea. Esta titánica labor se contextualiza en la era de la información digital caracterizada por la generación, distribución y conservación de una gran cantidad de recursos de información a través de las tecnologías de información y comunicación. Cada año la producción de información digital crece al doble. Se estima que en el año 2020 existan 44 zetabytes, es decir 44 trillones de gigabytes de información digital. (The digital Universe, 2014). En este contexto, miles de documentos sonoros, audiovisuales,



textuales, gráficos, circulan por Internet y día a día aparecen nuevos servicios agregados de información, productos, bancos de información y canales temáticos, entre otros. El 75% de la información digital es creada por individuos y empresas (The digital Universe, 2013).

La producción, distribución y conservación de documentos sonoros digitales, también forma parte de la información en la era digital. De ello da cuenta la creciente presencia de sitios como *eMusic*, *Allmusic*, *Spotify*, *Deezer*, *Muze*, *Amazon*, *Rhapsody* y *Lala*, en los cuales se pueden escuchar miles de documentos musicales de los más diversos géneros, a través de servicios de paga. Además, en las redes sociales, también hay sitios como *Soundhound*, *MySpace*, *Lastfm* y *Sound cloud*, que posibilitan a los usuarios publicar y escuchar materiales. También existen una amplia variedad de estaciones de radio en la red, así como la publicación sistemática de podcast y producciones sonoras que día a día se publican en internet.

Cada día se producen y transmiten miles de documentos sonoros de origen digital. De hecho, se ha afirmado que la web posee la mayor cantidad de grabaciones sonoras en formato digital de la historia (CLIR y LC, 2010). Gran parte de estas producciones digitales no se recopilan de forma sistemática y muchas se pierden irremediamente. De forma paradójica, ante la proliferación de contenidos digitales sonoros que se pueden escuchar, almacenar, intercambiar y producir en plataformas digitales, hay pocas garantías que estos documentos están siendo archivados y preservados de forma sistemática.

Esta situación explica que el investigador norteamericano Bill Klinger señale que si sólo el 17% de los cilindros grabados en un siglo en Estados Unidos pudieron ser recuperados; es muy probable que en 2110 un porcentaje similar de las producciones digitales realizadas en los últimos veinte años sobreviva. (Brylawski, Sam y Bamberger, Rob. 2010). Por ello, han comenzado a discutirse recomendaciones para establecer las mejores prácticas para la administración de colecciones sonoras de origen digital. Brylawski (2011) ha advertido del crecimiento de las colecciones sonoras digitales resultado de la reedición de discos de música, la aparición de servicios de audio a la carta y la producción de radios por internet.

En consecuencia, la acumulación de documentos digitales que se están produciendo diariamente define el tercer desafío a considerar en torno a la preservación digital sonora. En este sentido, investigaciones recientes señalan que el 89% de las instituciones europeas dedicadas a la preservación de la memoria audiovisual preservan documentos que son de origen digital. Es decir, que no tienen un equivalente en formato analógico o bien que no han sido digitalizados de un soporte analógico. (Stroeker, Natasha and Vogels, René, 2012).

Lo que significa que además de la colosal tarea para transferir contenidos en soportes analógicos a plataformas digitales que están en riesgo de perderse sino se digitalizan en los próximos años; los archivos sonoros tienen ante sí, el desafío de comprender la naturaleza digital de los nuevos documentos sonoros y al mismo tiempo incorporar en las estrategias de preservación digital sonora los miles de documentos sonoros de origen digital que circulan en las redes de información.



Cuarto desafío: La sustentabilidad del archivo digital sonoro.

La preservación digital sonora ha de garantizar la conservación y acceso a largo plazo de los contenidos de documentos sonoros grabados en diversos soportes, que hayan sido transferidos a plataformas digitales; y además, de documentos sonoros que son de origen digital. De lo cual se deriva que los archivos sonoros en plataformas digitales, se suman a la información digital como recurso vital en la economía del conocimiento y son fundamentales para la ciencia, la educación, la industria y el patrimonio cultural (Blue Ribbon, 2010). Por ello, el desarrollo y crecimiento de la industria del almacenamiento de datos digitales y la preservación digital constituyen ámbitos de interés político, como se expresa en *Conclusions on digital heritage as a strategic resource for a sustainable Europa*, documento en el que se reconoce que la herencia tangible, intangible y de origen digital posee un gran valor para la sociedad, desde el punto de vista del entorno social, económico y cultural y por lo tanto su gestión sustentable constituye un cambio estratégico en el siglo XXI (UE, 2014), de ahí que se emita el llamado para generar modelos políticos basados en la sociedad y el ciudadano para la preservación de la herencia cultural a largo plazo. (UE, 2014). El término sustentable no es reciente. De acuerdo con Bradely (2007) en los años 60 la palabra sustentabilidad comenzó a utilizarse con un sentido económico y en los años 80 se asoció con temas de medio ambiente.

En el entorno digital, a principios del siglo XXI se empezó a desarrollar un debate en relación con la sustentabilidad, probablemente derivado de la necesidad de dar respuesta a la pregunta formulada en el tránsito del mundo analógico al digital: ¿existe un soporte digital permanente para los documentos sonoros?. Esta pregunta expresa una de las preocupaciones derivadas de la transferencia de contenidos analógicos a plataformas digitales: la obsolescencia, como condición permanente de los archivos digitales.

Hasta ahora no se cuenta con una plataforma única de preservación digital sonora. Schuller y Teruggi (2006) han coincidido en señalar que la migración de información hacia nuevos sistemas de almacenamiento, una vez que los sistemas sean técnicamente obsoletos, es uno de los aspectos clave para una estrategia exitosa de preservación digital. La IASA (2006) ha establecido al respecto que dado que el hardware, el software y los sistemas son herramientas a disposición de un archivo, la planeación a largo plazo consiste en que los sistemas puedan ser cambiados sin modificar el contenido. La migración es una tarea cíclica que forma parte de la visión a largo plazo de un archivo digital sonoro, es decir con la perspectiva de un repositorio digital sustentable.

Un repositorio digital sustentable debe ofrecer la capacidad de conservar y recuperar objetos digitales a largo plazo, independientemente de la estructura organizativa y los recursos económicos del repositorio digital. Por ello, lo sustentable también hace alusión a estándares, arquitecturas y esquemas de metadatos que a su vez deben ser sustentables, abiertos y documentados para apoyar los nuevos sistemas de información que surjan a largo plazo. Bradley (2007) ha dicho que el enfoque sustentable contemporáneo se distingue por la concentración en las arquitecturas de



sistemas y esquemas que ayudarán en la gestión futura de la información digital. Una adecuada preservación digital permitirá que un objeto digital pueda ser consultado (Bradley, 2006) en el futuro.

En consecuencia, el quinto desafío de la preservación digital sonora es la creación de archivos digitales sustentables, diseñados a partir de considerar que la continuidad de la información digital va más allá del almacenamiento, del manejo de los datos y de su integración en el ciclo de vida del objeto de información. Un archivo digital sustentable debe incluir consideraciones técnicas, sociales y económicas. Por lo tanto un repositorio sonoro digital sustentable y un plan de preservación digital deben centrar su interés en la permanencia de los contenidos a través del tiempo.

CONCLUSIONES

La preservación digital de los archivos sonoros que constituyen el patrimonio intangible, expresado en sonidos, a través del cual podremos conocer la historia reciente y dejar testimonio de lo que hemos sido como sociedad, enfrenta cuatro desafíos que deberá afrontar:

Los estudios destinados a conocer cuantitativa y cualitativamente los documentos sonoros, constituyen la base para el desarrollo de proyectos y estrategias de digitalización y de creación de plataformas de preservación digital de colecciones sonoras. La preservación digital sonora debe basarse en el diseño de un plan a largo plazo, que se fundamente en información fidedigna de la cantidad y condiciones de los archivos sonoros. Por ello, el conocimiento de la cantidad, contenidos y condiciones de conservación de los documentos sonoros, constituye el primer desafío de la preservación digital de archivos sonoros.

El segundo desafío fue formulado como consecuencia del deterioro progresivo a que están expuestos los soportes sonoros. Día a día, se incrementa el riesgo de pérdida de miles de documentos que forman parte de nuestra identidad, dan cuenta de nuestra historia, son una muestra de la expresión y la creación científica, artística y cultural, son el testimonio de lo que hemos sido como sociedad. La experiencia y conocimiento acumulado en las dos últimas décadas debe ser apreciado y utilizado para afrontar la carrera contra el tiempo que borre nuestra herencia sonora. Por ello, la cooperación y colaboración sin fronteras, institucionales, geográficas y sociales, constituye una alternativa ante la extinción, por deterioro y pérdida de los sonidos que durante más de un siglo y medio han sido grabados en diversos soportes.

Se ha señalado que la digitalización es una tarea monumental, porque implica la transferencia de millones de contenidos sonoros en diversos soportes a plataformas digitales, por ello, el tercer desafío que se deberá afrontar es considerar que día a día se producen, distribuyen y almacenan miles de documentos sonoros cuyo origen es digital. Comprender la naturaleza del documento digital sonoro e incorporarlo a los planes y plataformas de preservación digital sonora, es otro de los más importantes aspectos a considerar si no queremos que una parte de la memoria expresada en sonidos no vuelva a ser escuchada.



El cuarto desafío es comprender que para garantizar la preservación a largo plazo, es necesario diseñar planes de preservación digital sonora sustentable, que se fundamenten en la permanencia de los contenidos a lo largo del tiempo. En consecuencia, la sustentabilidad del archivo digital sonoro es el más grande desafío de la preservación digital.

REFERENCIAS

Addis, M., Allasia, W., Bailer, W., Boch, L., Gallo, F., & Wright, R. (2010). "100 million hours of audiovisual content: Digital preservation and access in the Presto Prime project". En *Proceedings of the 1st International Digital Preservation Interoperability Framework Symposium* (p. 3). ACM.

ARSC Technical Committee (2011). *Study of Embedded Metadata Support in Audio Recording Software*. Association of Recorded Sound Collections. Disponible en: www.arsc-audio.org/pdf/ARSC_TC_MD_Study.pdf (4 de agosto de 2014).

IASA (2003). *TC03. La salvaguardia del patrimonio sonoro: ética, principios y estrategias de preservación*. Asociación Internacional de Archivos Sonoros y Audiovisuales (IASA)-Radio Educación, México.

_____ (2006). *Lineamientos para la producción y preservación de objetos de audio digitales*. TC-04. IASA-Radio Educación, México.

Blue Ribbon (2008). *The Blue Ribbon Task Force on Sustainable Digital Preservation and Access. Interim Report*. Disponible en: http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf 1 de agosto de 2014.

Boston, G. (2003). "Survey of Endangered Audiovisual Carriers." IASA y UNESCO, París. Disponible en: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/programme_doc_survey_report.pdf (29 de julio de 2014).

Bradley, K. (2006). "Digital sustainability and digital repositories," en DSpace on 2011-01-04T05: 34: 33Z (GMT).

_____ (2007). "Defining digital sustainability", en *Library Trends*, 56(1), 148-163.

Brylawski, Sam y Bamberger, Rob (2010). *The State of Recorded Sound Preservation in the United States: A national Legacy at Risk in the Digital Age*. Council on Library and Information Resources at the Library Congress, Washington, D.C. 169 p.



CLIR y CL (2010). *The state of recorded sound preservation in the United States: A National Legacy at Risk in the Digital Age*. Council on Library and Information Resources and the Library of Congress. Washington, D.C.

Edmondson, R. (2010). "Philosophical fundaments in the digital age". *Conferencia presentada en el Cuarto Seminario Internacional de Archivos Sonoros y Audiovisuales*. CONACULTA, México.

Hoog, Emmanuel (2005). "¿Guardar todo?. *Los dilemas de la memoria en la edad mediática*". Radio Educación, México.

IFEA (2003). *Comité de Censo y Valoración de Registros Sonoros y Audiovisuales Etnográficos de los Países Andinos*. Organizado por el Instituto Francés de Estudios Andinos en la Ciudad de Bogotá los días 1, 2 y 3 de 2003. Disponible en: http://www.ifeanet.org/multimedia/comite/doc/Seminario-2003_12_01-Conclusiones.pdf (16 de mayo de 2014).

Klijn, E. & Lusenet, Y. D. (2008). *Tracking the reel world*. European Commission on Preservation and Access.

Rodríguez, Perla (2012). *El archivo sonoro. Fundamentos para la creación de una fonoteca nacional*. Library Outsourcing.

Schüller, Dietrich (2006). "La preservación de la herencia digital", en Rodríguez, Perla (Comp.). *Memorias del Tercer Seminario Internacional. La Preservación de la memoria audiovisual en la sociedad digital*. Radio Educación, México.

_____ (2008). *Audiovisual research collections and their preservation*. European Commission on Preservation and Access. The digital universe. Disponible en <http://www.emc.com/leadership/programs/digital-universe.htm>, (4 de diciembre de 2013).

_____ (2012). "The Role of Digitization in the Preservation of Audiovisual Documents", en *Preservation of Digital Information in the Information Society. Proceedings of the International Conference, Moscow*. Disponible en http://www.ifapcom.ru/files/News/Images/2013/dig_pres_is_eng.pdf (30 de marzo de 2014).

_____ (1997). *Strategies for the safeguarding of audio and video materials in the long term*. United Nations Educational Scientific and Cultural Organization Publications -CII WS, 292-297.



Stroeker, Natasha and Vogels, René (2012). *Survey Report on Digitisation in European Cultural Heritage Institutions 2012*, EUMERATE Thematic Network, Panteia (NL). Disponible en <http://www.enumerate.eu/fileadmin/ENUMERATE/documents/ENUMERATE-Digitisation-Survey-2012.pdf> (10 de diciembre de 2013).

Teruggi, Daniel (2006). "Presto Space Factories: Proyecto europeo de investigación sobre la conservación y digitalización de los archivos audiovisuales" en *Rodríguez, Perla (Comp.). Memorias del Tercer Seminario Internacional. La Preservación de la memoria audiovisual en la sociedad digital*. Radio Educación. México.

The digital universe (2013). Disponible en <http://www.emc.com/collateral/analyst-reports/idc-digital-universe-united-states.pdf>, (1 de agosto de 2014).

UE (2014). *Conclusions on cultural heritage as a strategic resource for a sustainable Europe*. Disponible en http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/educ/142705.pdf (25 de mayo de 2014).

UNESCO(1980). *Recomendación sobre la salvaguarda y la conservación de las imágenes en movimiento*. Conferencia General de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. Disponible en http://portal.unesco.org/es/ev.phpURL_ID=13139&URL_DO=DO_PRINTPAGE&URL_SECTION=201.html (3 de octubre de 2009).

UNESCO-UBC (2012). *Declaración de Vancouver. La memoria del Mundo en la era digital: digitalización y preservación*.

Wright, Richard. (2004). *Digital preservation of audio, video and film*. *Vine*, 34(2), 71–76. doi:10.1108/03055720410550869

_____ (2011). *Audiovisual digital preservation*. Status report 4. ICT-2007-3-231161. Disponible en http://preservationguide.co.uk/wp-content/uploads/2013/11/PP_WP7_D7.1.5_Annual_AV_Status_2011.pdf (1 de agosto de 2014).

_____ (2011) "How can invisible files stored somewhere on masstorage-perhaps even in the cloud.ever claim authenticuty?", en *Zorgen Voor onzichtbare assets. Over Het Behoud Van Digitale AV-Collectties*. Hilversum.



_____ (2012). *Preserving Moving Pictures and Sound*. Digital Preservation Coalition, Gran Bretaña.

_____ (2012). *Moving pictures and sound*. DPC Technology Watch 12-01. Disponible en <http://dx.doi.org/10.7207/twr12-01> (febrero de 2014.).



EVALUACIÓN AUTOMÁTICA DE PRESERVACIÓN MEDIANTE LA UTILIZACIÓN DE TAREAS DE CURACIÓN EN DSPACE

Terruzzi Franco Agustín¹, Lira Ariel Jorge², De Giusti Marisa Raquel³, Villarreal Gonzalo Lujan⁴

1. Estudiante avanzado de Licenciatura en Sistemas en la UNLP. Servicio de Difusión de la Creación Intelectual, Proyecto de Enlace de Bibliotecas, Universidad Nacional de La Plata (UNLP).
2. Licenciado en Informática. Servicio de Difusión de la Creación Intelectual, Proyecto de Enlace de Bibliotecas, Universidad Nacional de La Plata (UNLP).
3. Ingeniera en Telecomunicaciones y Profesora de Letras. Servicio de Difusión de la Creación Intelectual, Proyecto de Enlace de Bibliotecas, Universidad Nacional de La Plata (UNLP). Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA).
4. Doctor en Ciencias Informáticas. Servicio de Difusión de la Creación Intelectual, Proyecto de Enlace de Bibliotecas, Universidad Nacional de La Plata (UNLP).

INTRODUCCIÓN

Muchos repositorios institucionales asignan parte de sus recursos a las actividades ligadas a la difusión y la preservación [1]. La difusión se realiza mediante una serie de acciones cuyo objetivo es lograr que los elementos del repositorio alcancen a la comunidad; la preservación se refiere a la persistencia en el tiempo de tales elementos, a fin de asegurar su acceso y usabilidad a largo plazo. Ambas actividades están influenciadas por los aspectos políticos, económicos y tecnológicos que rigen al repositorio. Las actividades de preservación están enmarcadas por las políticas del repositorio: el desarrollo de tales políticas deben ser el primer paso para garantizar las acciones de preservación [2]. Estas acciones implican generalmente la validación de los elementos del repositorio a partir de las restricciones indicadas en sus políticas, lo que puede realizarse tanto forma manual como con la ayuda de alguna herramienta que permite su automatización [3]. Típicamente, durante la evaluación de un recurso se toman algunas de sus propiedades y se las compara con parámetros esperados según se ha establecido previamente. Estos valores esperados surgen, en muchos casos, a partir de modelos estándares como la norma OAIS, lo que permite definir la generación de un plan de preservación basado en un contexto normalizado y acotado. No obstante, los cambios en las políticas y la evolución tecnológica constante hacen que estas validaciones deban revisarse permanentemente, dificultando la ejecución de las tareas de evaluación automáticas o semiautomáticas [4].



Por otro lado, las tareas de evaluación y control pueden efectuarse manualmente mediante comprobaciones realizadas desde el software en uso o incluso ejecutando consultas a la base de datos, y por medio del establecimiento de un flujo de trabajo para la gestión de los cambios sobre los elementos. Sin embargo, si bien esto permite alcanzar un alto nivel de precisión al evaluar los elementos del repositorio, esta labor demanda mucho tiempo y, como todo proceso manual e iterativo, suele ser propenso a fallos.

El SEDICI [5] es un caso típico de repositorio digital donde se presenta la situación descrita en los párrafos anteriores. Desde su creación en el año 2003, el repositorio institucional de la Universidad Nacional de La Plata [6] ha evolucionado tanto en las políticas como las tecnologías que utiliza [7]. En la actualidad reúne cerca de 35000 obras que provienen de toda la UNLP y con tipologías muy variadas. Este crecimiento sostenido y evolución continua a lo largo de los años demanda grandes esfuerzos para realizar los controles y evaluaciones sobre los ítems, así como también actualizar los mecanismos para realizar dichos controles.

En este contexto, el presente trabajo propone una herramienta que permite evaluar e informar el “*estado actual de preservación*” de los elementos del repositorio de manera automática. La herramienta permite personalizar tanto los controles como las actualizaciones a realizar, y dispone de mecanismos para seleccionar los elementos del repositorio en donde aplicar dichos controles. Esto se enmarcó en el desarrollo de Dspace [8], para que se pueda compartir con la comunidad especializada en el tema.

DISEÑO

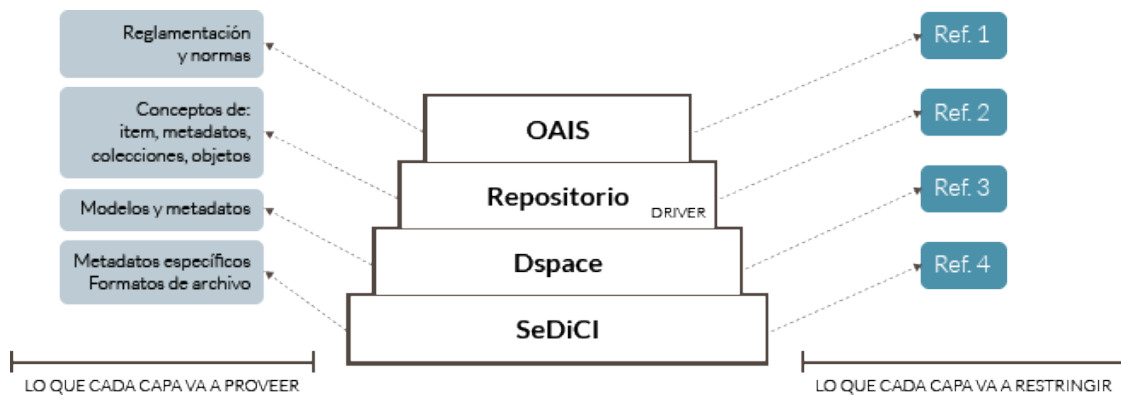
Se estableció un diseño que permite representar la variabilidad de las reglas de preservación según su alcance, y se implementó un método para construir nuevas reglas y ejecutarlas sobre elementos del repositorio especificados previamente. Cada regla a ejecutar forma parte de un contexto determinado, que define su nivel de abstracción así como también las reglas de validación aplicables a dicho contexto. En este sentido, se pensaron diferentes *niveles de especificación*¹ para la definición de las reglas: la norma OAIS como nivel más general, el nivel de Repositorio para definir elementos y restricciones típicas de los repositorios, el nivel Dspace donde se determinan aspectos propios de este software, y finalmente el nivel SEDICI que toma elementos definidos particularmente para este repositorio (ej. metadatos propios). Las características de cada nivel definen la reusabilidad de las reglas, su aplicabilidad (desde las más abstractas hacia las más específicas) y su relevancia. Para cada nivel se determinan 1) las restricciones a evaluar, 2) las propiedades a verificarse y 3) el contexto del cual se desprenden. Lógicamente, a medida que el nivel se vuelve más específico, también lo hacen sus restricciones a partir del repositorio donde se aplican las reglas y, al mismo tiempo, su implementación se vuelve más simple y factible.

1

Utilizaremos este concepto en el trabajo, para referirnos al contexto de aplicación de la regla que se describe.

Esta variabilidad hace que las reglas de mayor abstracción sean aplicables a la mayoría de los repositorios digitales, pero también sean más costosas en tiempo y esfuerzo en el momento de su implementación.

En la Imagen 1 se muestran los niveles de especificación establecidos y los diferentes dominios y restricciones que presentan.



<p>Referencia 1</p> <ul style="list-style-type: none"> - Fixity (se va a evaluar de acuerdo a la documentación planteada) - Autenticidad (se acuerdo a la forma propuesta por OAIS) - Contexto (de acuerdo a la forma propuesta por OAIS) - Accesibilidad (lo necesario para asegurar que siempre se pueda acceder al ítem) al momento evaluado en relación a la representación de los objetos del repositorio. - Legibilidad (lo necesario para asegurar que los formatos de ítems pueden leerse y reconocerse) al momento evaluado en relación a la representación de los objetos del repositorio.
<p>Referencia 2</p> <p>Se debe tener en cuenta que deben cumplirse ciertas políticas, tales como:</p> <ul style="list-style-type: none"> - Reglamentación de Metadatos: p.e. Directivas Driver 2 u OpenAire - Contenido: tipologías y formatos permitidos - Licencias: Licencias requeridas con las que debe contar el ítem, p.e. Licencia de distribución y uso. - Preservación: Las políticas de preservación engloban en la fase tecnológica: La conservación del archivo original, la conservación en algún formato apto para la preservación. La existencia de un identificador persistente que haga ubicuo al objeto y permita su acceso "para siempre". Reglas de Accesibilidad (que pueda verse siempre) esto significa incluso que se realice la migración del objeto para mantenerlo en el tiempo. Validaciones sobre el objeto digital.
<p>Referencia 3</p> <p>Las restricciones impuestas para DSpace :</p> <p>Cálculo y validación del checksum con MD5. Handle -> (siempre existe en DSpace), validación (que sea válido), que no sea el que trae la instalación puesto por default. Formatos-> que los formatos de archivo sean los que DSpace considera válidos</p>
<p>Referencia 4</p> <p>Las restricciones impuestas por SEDICI están vinculadas fundamentalmente a las políticas del repositorio: políticas de contenido, de preservación, de metadatos, etcétera.</p>

Figura 1: Niveles de especificación definidos: restricciones y propiedades de cada uno. Mientras más específico es el nivel, las propiedades y restricciones que ofrece se tornan más dependientes de la implementación y del repositorio en el que se está trabajando.

Se trabajó en una primera iteración sobre la aplicación de reglas aisladas sobre conjuntos de ítems, y a posteriori sobre la posibilidad de añadir nuevas reglas, quitarlas y modificarlas. Por otro lado, se estableció un nuevo criterio para evaluar el *grado de preservación*² de un determinado recurso, lo que permite obtener una medida uniforme de comparación entre los elementos. Para esto la validación de cada regla retorna un valor numérico, que representa el grado con el que el elemento cumple con dicha regla. Dado que este nivel de adecuación a una regla puede variar de acuerdo con el contexto en que se la aplica, se incorporó un mecanismo que permite definir la importancia de cada regla junto con el orden y contexto de aplicación. De este modo, es posible preparar “*recetas*”: conjuntos de reglas que poseen un peso y un orden de ejecución que depende de la receta, y que se calculan en el momento de ejecutar el validador. Este validador fue diseñado sobre un modelo de objetos que hace énfasis en el modelado de reglas y en la posibilidad de configurar el modo de aplicación de las mismas.

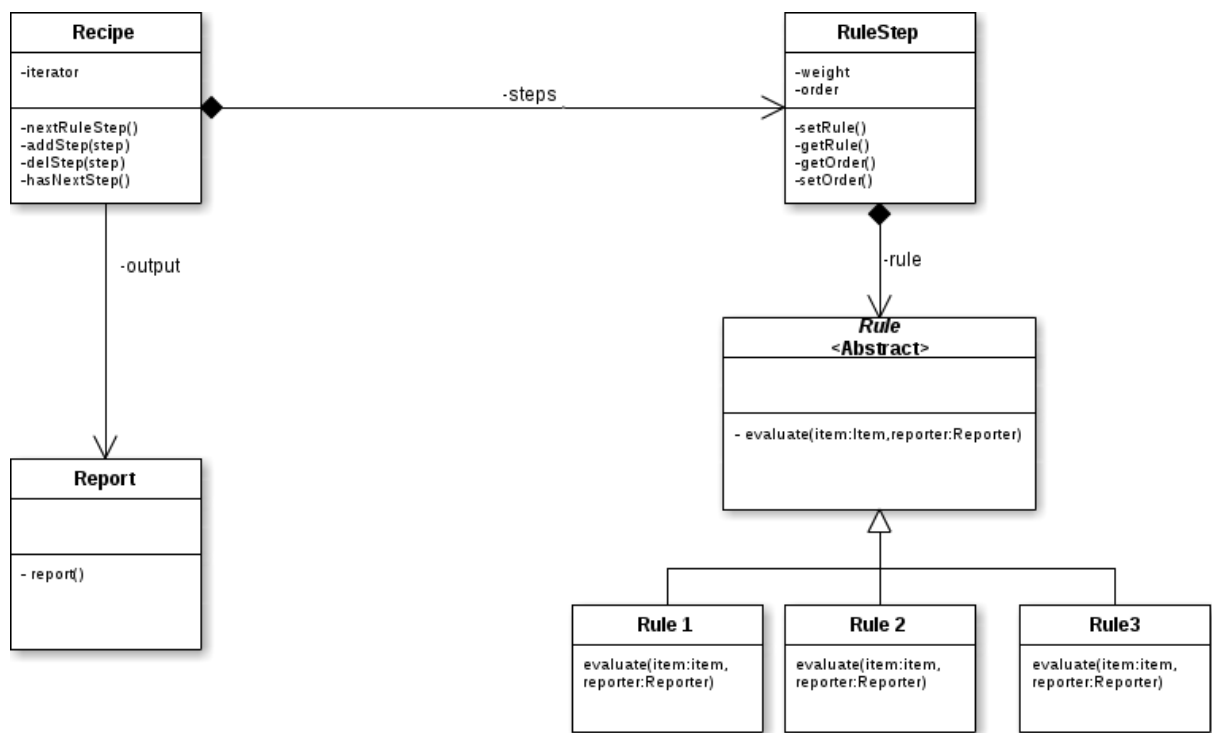


Figura 2: Diagrama UML del validador propuesto. Las reglas (Rule) conforman una jerarquía fácilmente extensible, y se organizan en etapas con un orden y un peso (RuleStep). Estas etapas componen recetas (Recipe), cuyo resultado es reflejado en un reporte (Report).

2

El grado de preservación es un concepto utilizado en este trabajo para establecer una medida cuantitativa de la posibilidad de preservar un determinado ítem o elemento en un repositorio.



Se estableció así un mecanismo normalizado de definición de reglas, en donde la administración del repositorio puede definir una regla fácilmente, a partir de una definición semántica de la misma. En el SEDICI, se concilió³ una estandarización para la especificación de las reglas de validación, que se escriben en base a una estructura como se muestra en el ejemplo a continuación:

Enfoque: DSpace

Nombre: “Verificar validez de Handle”

Restricciones: Handle siempre existe en DSpace, puede ser el handle predefinido

Regla: El ítem contiene un handle válido y no se trata del Handle por defecto (123456789)

Metadato(s) Asociado: dc.identifier.uri (sedici.identifier.handle)

Respuesta esperada: True => Válido

False => Inválido

Cada regla se implementa mediante una clase JAVA [9] como subclase de la clase abstracta Rule. El uso del patrón de diseño *Template Method* [10] simplifica la generación de nuevas reglas, requiriendo la implementación de la regla de evaluación (método *evaluate*). El resultado de la ejecución de cada receta, con sus etapas y reglas, es enviado hacia un reporte. La clase Report permite encapsular la información que se desea generar en el reporte, alterar el grado de granularidad (nivel de especificación) de la misma, e incluso adaptar el formato de salida según los requerimientos del usuario (archivo de texto plano, csv, salida en pantalla, documento HTML, etc.)

IMPLEMENTACIÓN

El diseño antes propuesto debía acoplarse al contexto de SEDICI, principalmente al software que sustenta este repositorio. Por este motivo se buscó adaptar el modelo hacia una estructura compatible con un módulo de Dspace. Esto permite reutilizar todo el modelo de datos y de objetos del repositorio y, de este modo, es posible implementar rápidamente una solución orientada a la reutilización.

Dspace ofrece la posibilidad de desarrollar módulos tradicionales, pero también permite gestionar tareas de curación [11] [12], lo que resulta de suma utilidad para esta solución. Se diseñó entonces una *tarea de curación* específica capaz de ejecutar la validación correspondiente y que conozca cómo y cuáles reglas aplicar. Esto llevó a que en una segunda iteración, se construya el diagrama UML [13], mostrado en la imagen 2.

3

Se estableció de común acuerdo entre el personal técnico, administrativo y los expertos la temática.

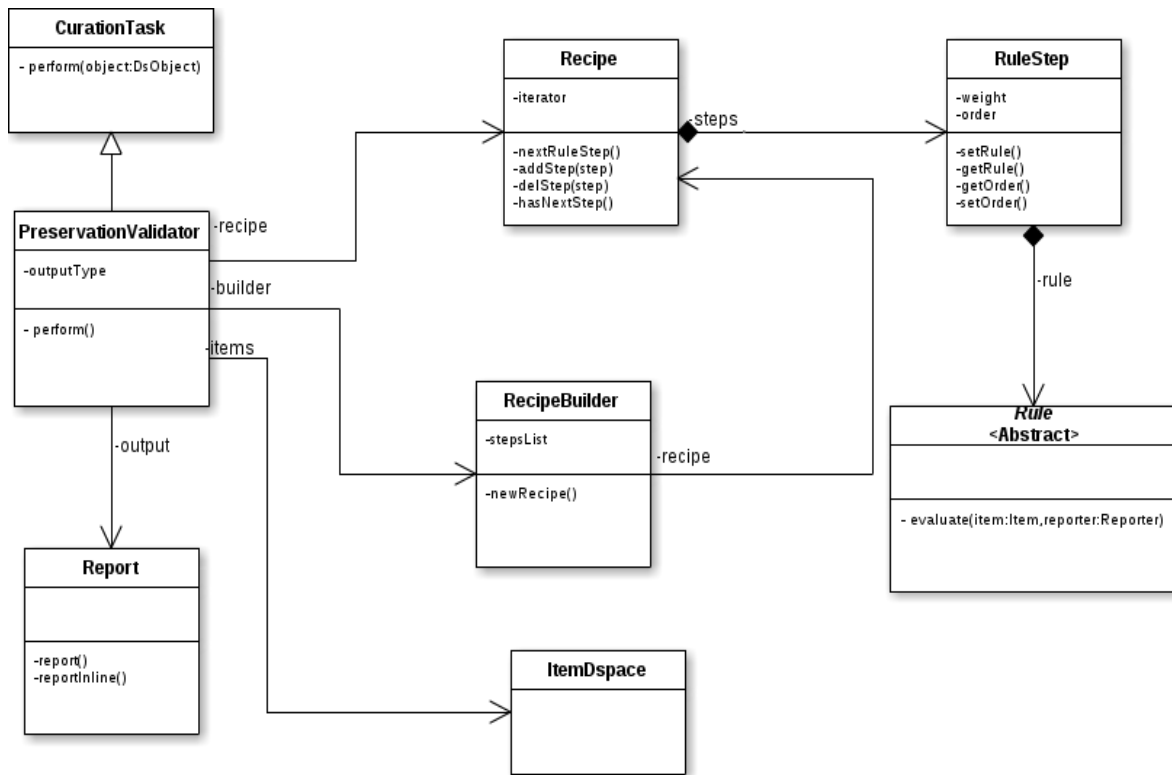


Figura 3: Diagrama UML del validador propuesto a partir de la adecuación de la solución a una tarea de curation en Dspace. El validador es en sí mismo una tarea de curación y se utiliza como tal. Con cada ejecución, aplica una serie de reglas (Rules) ordenadas en recetas de diferente tipo, a un conjunto específico de elementos del repositorio (ItemDspace).

A partir de este diseño, se implementaron las clases JAVA correspondientes para que la tarea de curación pudiera ejecutarse. Un requisito importante de este desarrollo es poder incorporar y modificar recetas, basadas en reglas preexistentes, sin la necesidad de modificar el código fuente de la aplicación (lo que llevaría a recompilar, *deployment*, etc). Se buscó entonces un método de definición de recetas, que detecta las configuraciones en tiempo de ejecución. Esto permite intercambiar las recetas, reglas y etapas de manera muy simple y sin necesidad de conocer la implementación del validador. Para lograr esto, el desarrollo se basó en las especificaciones de las tareas de curación en Dspace, que permiten definir y cargar una serie de *propiedades* para cada tarea, un método por el cual una tarea puede conocer un valor de una configuración en un archivo determinado en tiempo de ejecución.

Se definió una clase *builder* [14] que es ejecutado por la tarea de curación y que instancia las recetas con las propiedades definidas. Estas propiedades se cargan en tiempo de ejecución, y son leídas desde el archivo de configuración específico del validador, que se encuentra con el nombre de *evaluatePreservation.cfg*, con la estructura que se muestra en el ejemplo a continuación:



```
#-----#  
#-----RULE VALIDATOR Curation Task Configurations-----#  
#-----#  
#----- Configuration properties used solely-----#  
#-----by the rule validator task -----#  
#-----#  
#-----@author Franco Agustín Terruzzi-----#  
validator.rules=\  
ar.edu.unlp.sedici.dspace.curation.preservationHierarchy.preservationRules.HandleValidationRule = r1,\  
ar.edu.unlp.sedici.dspace.curation.preservationHierarchy.preservationRules.Md5ValidationRule = r2,\  
ar.edu.unlp.sedici.dspace.curation.preservationHierarchy.preservationRules.LicenceValidationRule = r3,\  
ar.edu.unlp.sedici.dspace.curation.preservationHierarchy.preservationRules.ExistMetadataValidationRule = r4  
  
#rules properties  
#rules properties must be written starting with rulename.propertyName and must be used like Strings  
#the rules properties values depends directly of the recipes configurations, hence it must be specified in the recipe  
configuration file  
# example  
    # r1.predefined.Handle=1234567/8  
# see the recipe configuration file for more details  
  
#recipes must be written in the next format: recipe.rules = rule1, weigth1; rule2, weigth2; rule3, weigth3  
#recipes must be written in a distinct cfg file see HandleValidation.cfg  
recipe.name = Empty Recipe  
recipe.rules= []
```

Ejemplo 1: Configuración típica del archivo evaluatePreservation.cfg. Las tareas de curación en Dspace permiten configurar ciertas propiedades específicas de una tarea, mediante la formulación de un archivo de texto plano de configuración que posea el mismo nombre que la tarea.

La información especificada en el archivo de configuración permite al validador definir qué reglas aplicar, en qué orden aplicarlas y establecer el peso para cada una. Para permitir la aplicación de distintas recetas, se utilizó un *sistema de herencia* [15] de tareas de curación, en donde se define en el archivo de configuración correspondiente no sólo la tarea de curación en sí sino también la receta que dicha tarea aplica. Estos archivos de configuración fueron denominados *archivos de configuración de recetas* y deben ser definidos por los diseñadores de las reglas y de las recetas al momento de ejecutarlas. Esto permite, adicionalmente, que la ejecución del validador con una nueva receta sólo requiera añadir una nueva entrada en el archivo de configuración correspondiente que haga referencia al *archivo de configuración* específico de la nueva receta (*curate.cfg*). Esto se visualiza en el siguiente ejemplo:



```
#-----#  
#-----CURATION SYSTEM CONFIGURATIONS-----#  
#-----#  
# Configuration properties used solely by the curation system #  
#-----#  
  
### Task Class implementations  
  
### Task Class implementations  
  
plugin.named.org.dspace.curate.CurationTask = \  
  org.dspace.ctask.general.NoOpCurationTask = noop, \  
  org.dspace.ctask.general.ProfileFormats = profileformats, \  
  org.dspace.ctask.general.RequiredMetadata = requiredmetadata, \  
  org.dspace.ctask.general.ClamScan = vscan, \  
  org.dspace.ctask.general.MicrosoftTranslator = translate, \  
  org.dspace.ctask.general.MetadataValueLinkChecker = checklinks, \  
  ar.edu.unlp.sedici.dspace.curation.LinkControlledMetadata = linkMetadata, \  
  ar.edu.unlp.sedici.dspace.curation.CopyMetadata = copyMetadata, \  
  ar.edu.unlp.sedici.dspace.curation.FixControlledMetadata = fixControlledMetadata, \  
  ar.edu.unlp.sedici.dspace.curation.EvaluatePreservation = evaluatePreservation, \  
  ar.edu.unlp.sedici.dspace.curation.EvaluatePreservation = evaluatePreservation.HandleValidation, \  
  ar.edu.unlp.sedici.dspace.curation.EvaluatePreservation = evaluatePreservation.Md5Validation, \  
  ar.edu.unlp.sedici.dspace.curation.EvaluatePreservation = evaluatePreservation.ExistMetadataValidation, \  
  ar.edu.unlp.sedici.dspace.curation.EvaluatePreservation = evaluatePreservation.LicenceValidation  
  
# add new tasks here
```

Ejemplo 2: configuración definida en el archivo curate.cfg. Para añadir una tarea de curación nueva, debe añadirse una línea en este archivo de configuración de Dspace. Cada receta que se quiera aplicar y configurar en tiempo de ejecución, debe añadirse en este archivo como una nueva tarea de curación.

En una instalación normal de Dspace, los archivos de configuración se encuentran en <directorio_instalación>/config/modules/NombreCuration.cfg.

De la misma forma se pueden definir un conjunto de propiedades de cada regla en una receta en específico. Esto surgió de la idea de que una regla puede variar levemente desde un contexto de aplicación a otro (por ejemplo, yendo de un ámbito más general a uno más específico), sin la necesidad de que tenga que implementarse por completo. Por este motivo en el archivo de configuración que representa a la receta correspondiente, es posible agregar algunas propiedades que estarán disponibles para las reglas durante su ejecución. Para esto, se utiliza la siguiente notación:



```
#-----#  
#-----RULE VALIDATOR Curation Task Configurations-----#  
#-----#  
#----- Configuration properties used solely-----#  
#-----by the rule validator task -----#  
#-----use this configuration to set the choosen recipe name-#  
#-----@author Franco Agustín Terruzzi-----#  
recipes.name = ExistMetadataValidation  
recipes.rules = r4,1;
```

#metadata to validate in this recipe

```
r4.metadata.validate=\  
dc.description.provenance,\  
dc.identifier.uri,\  
sedici.identifier.uri,\  
mods.location
```

Ejemplo 3: notación de las propiedades de una regla de validación en el archivo evaluatePreservation.ExistMetadataValidation.cfg. Las propiedades para una regla inician con el nombre de la misma y siguen el formato mostrado NombreDeRegla.Propiedad = valor1, valor2,valor3 (para valores con saltos de línea se debe agregar \).

El desarrollo aquí descrito permite ejecutar una tarea de curación que valide una serie de ítems, lo que puede resolverse directamente desde un intérprete de comandos y con solo un comando:

```
./dspace curate -r -t evaluatePreservation.<RecetaAValidar> -i <handleItem | handleComunidad |  
handlColección | all> -v
```

Donde:

- -r define el tipo de reporte para gestionar la salida
- -l indica que la salida será la estándar
- -t permite definir el tipo de tarea que se va a ejecutar
- -i permite definir los tipos de ítems en donde se aplicará la tarea
- -v hace más comunicativa la salida de la tarea

A partir de estas primeras reglas se pudieron definir otras que se desprendían de los *niveles de especificación* ya mencionados, y se aplicaron recetas compuestas por varias reglas, con configuraciones, órdenes y pesos distintos. En este contexto, se construyó una receta que, para cada ítem, verifica una serie de propiedades que representan su *grado de preservación* con respecto a lo establecido como *Información Descriptiva de Preservación (PDI)* [16]. Se establecieron reglas para verificar:

- **Procedencia:** la información de procedencia es información que Dspace completa de manera predeterminada para cada ítem al momento de instanciarlo. Se comprueba que esta información exista y sea válida. Para esto se consulta el metadato estándar de dublin-core *dc.description.provenance* [17].

- **Contexto:** se verifica el contexto externo del ítem, validando que la información de contexto exista y esté asociada correctamente al ítem. Para hacer esto, se consultan los metadatos correspondientes a la localización física (mods.location) y electrónica (sedici.identifier.uri) asignados al ítem y verifica su existencia y validez.
- **Referencia:** Se verifica a través de la existencia de identificadores persistentes asignados. En este caso, la regla valida que el Handle que fue asignado para el ítem y que figura en el metadato dc.identifier.uri sea válido. Se considera a un Handle como válido cuando no es el que viene predefinido por Dspace (12345678) y cuando el contenido alojado en el metadator dc.identifier.uri es igual al guardado en la tabla Handle de la base de datos de Dspace.
- **Integridad:** se evalúa utilizando el checksum, el cual es calculado al momento de guardar el ítem y volviendo a calcularlo durante la validación del mismo. En este caso, se verifica que para cada *bitstream* de cada *bundle* [18] del ítem correspondiente, se mantenga su integridad. Para esto recurre al módulo *checker* [19] que brinda Dspace, que se utiliza para verificar el checksum correspondiente a cada ítem, comparándolo contra un checksum esperado almacenado en la base de datos. El ítem sólo es válido cuando todos los bitstreams relacionados con él han arrojado "CHECKSUM_MATCH" como resultado del chequeo. Cabe mencionar que para el caso de Dspace, el checksum se verifica y calcula usando el algoritmo MD5 [20].
- **Derechos:** todo ítem debe tener una licencia asignada, por lo que el metadato correspondiente a la Licencia CC debe existir siempre. Además, la URI de la Licencia debe también existir y debe ser la correspondiente con la Licencia elegida. En SEDICI, esto se valida chequeando la existencia y veracidad de los metadatos sedici.license y sedici.rights.uri.

Con la definición de estas reglas se pudo aplicar una receta que verifique la PDI de cada elemento del repositorio, obteniendo un reporte automatizado del estado de preservación del repositorio en el momento de la ejecución.

EJECUCIÓN

Se utilizó la herramienta por primera vez, con una única regla que consistía en verificar si los elementos escogidos tenían o no un Handle [21], y si este Handle no era el predefinido para dspace (123456789).

En una primer prueba se ejecutó en un solo elemento, el 10915/10672⁴, el comando de consola para realizar esta tarea quedó con el formato siguiente:

```
./dspace curate -r - -t evaluatePreservation.HandleValidation -i 10915/10672 -v
```

Se obtuvo un resultado correcto, cuyo reporte se muestra a continuación:

```

/home/nesstor/workspace_indigo_2/dspace-sedici/install/bin : sudo
Archivo Editar Ver Marcadores Preferencias Ayuda
terru@pc-nesstor ~/Java/workspace/dspace-sedici/install/bin $ sudo ./dspace curate -r - -t evaluatePreservation.HandleValidatio
Adding task: evaluatePreservation.HandleValidation
Starting curation
Curating id: 10915/10672

Procesando Item con id: 9620
El Handle del item: 9620 es válido, el item fue evaluado con 1
La puntuación total del item procesado es: 1,0
El item: 9620 se finalizó de procesar con éxito
Ending curation. Elapsed time: 107
terru@pc-nesstor ~/Java/workspace/dspace-sedici/install/bin $
  
```

Figura 4 : consola con resultados de la ejecución para el Ítem 10915/10672. Puede verse que la validación fue exitosa debido a que se corrobora para el Ítem especificado, que el handle asignado es un handle válido.

4

En el SEDICI, así como en la mayoría de los repositorios que utilizan Dspace, los elementos se identifican de manera unívoca con el handle que tienen asignado. Este handle es un tipo de identificador persistente, es decir, que ya el hecho de utilizarlo es en sí mismo un mecanismo de preservación.



Se probó entonces, la ejecución para una colección particular completa, identificada en SEDICI con el Handle 10915/123 (<http://sedici.unlp.edu.ar/handle/10915/123>) y se aprovechó el redireccionamiento⁵ de la salida de la consola, para guardar el reporte en un archivo. Se obtuvo también un resultado satisfactorio de la evaluación. Se utilizó la siguiente sentencia:

```
./dspace curate -r - -t evaluatePreservation.HandleValidation -i 10915/123 -v > result.txt
```

Por último se ejecutó esta única regla para todo el repositorio, con el objetivo de evaluar su desempeño tanto en tiempo de ejecución y en utilización de memoria como en eficacia. Dado que el reporte era demasiado extenso, se lo almacenó nuevamente en un archivo en disco, para poder procesar posteriormente los resultados. La sentencia de ejecución se indica a continuación (observar el uso de la palabra clave "all" para identificar a todos los ítems):

```
./dspace curate -r - -t evaluatePreservation.HandleValidation -i all -v > result.txt
```

Los resultados para la validación del Handle en todos los elementos del repositorio fueron exitosos. La tarea de curación puede ejecutarse en un breve período de tiempo y utiliza la memoria de manera eficiente, esto se puede apreciar en la Figura 4, donde (aunque para un solo elemento) el tiempo estimado fue de sólo 107 ms y donde el proceso no generó una carga demasiado alta en el rendimiento general del sistema en donde se ejecuta.

CONCLUSIONES

En varias situaciones en donde se utilizó la herramienta sobre el repositorio SEDICI, se pudieron identificar elementos que no superaron alguna validación y, por consiguiente, implementar un flujo de trabajo para realizar las correcciones pertinentes sobre los elementos. Este mecanismo de automatización de tareas de control de calidad, ha tenido resultados satisfactorios no sólo a la hora de su utilización, sino también como herramienta para evaluar posibles situaciones de mejora en las políticas de preservación del repositorio.

La visualización amigable de los resultados, ha permitido a la administración realizar pruebas y validaciones personalizadas, incluso ha sido útil para realizar algunos experimentos de evaluación del estado de preservación del repositorio [22].

5

El redireccionamiento es un método para dirigir la salida estándar de la consola hacia un archivo, con el objetivo de que esta salida pueda guardarse para análisis posteriores. Existe mucha información en línea de esta técnica, puede verse una explicación bien fundada en : http://www.robvanderwoude.com/battech_redirection.php



Adicionalmente, este desarrollo promueve la generación rápida de nuevas reglas de evaluación gracias a su diseño estructural. Esta facilidad en la generación de reglas expuso rápidamente la posibilidad de realizar mejoras sobre la herramienta, con el objetivo de mejorar su interfaz de usuario. En este contexto ha surgido la idea de implementar un mecanismo que permita a cualquier administrador del repositorio definir y ejecutar estas reglas de validación sin la necesidad de contar con conocimientos técnicos sobre la herramienta. La definición de estas reglas podría hacerse incluso desde la interfaz de administración del repositorio y podría darse, por ejemplo, mediante un lenguaje simple de utilizar que aplique las validaciones correspondientes y devuelva los reportes adecuados.

REFERENCIAS

1. Sánchez, S., & Melero, R. (n.d.). La denominación y el contenido de los Repositorios Institucionales en Acceso Abierto : base teórica para la “Ruta Verde.” Revisado Julio, 16, 2014 desde <http://digital.csic.es/handle/10261/1487>
2. Yuan, L. (2011). Institutional Repositories and Digital Preservation: Assessing Current Practices at Research Libraries, *17*(5/6).doi:10.1045/may2011-yuanli
3. De Giusti, M. R., Lira, A. J., Villarreal, G. L., & Texier, J. (2012, November 1). Las actividades y el planeamiento de la preservación en un repositorio institucional. *BIREDIAL - Conferencia Internacional Acceso Abierto, Comunicación Científica Y Preservación Digital*. Revisado Julio, 16, 2014 desde <http://hdl.handle.net/10915/26045>
4. Hamburger, S. (2003). The State of Digital Preservation: An International Perspective. *Library Collections, Acquisitions, and Technical Services*, *27*(3), 373–374. doi:10.1080/14649055.2003.10765943
5. SEDICI - Repositorio de la Universidad Nacional de La Plata . Revisado Julio, 16, 2014 desde <http://sedici.unlp.edu.ar/>.
6. Universidad Nacional de La Plata (UNLP). Revisado Julio, 16, 2014 desde <http://unlp.edu.ar/>
7. De Giusti, M. R. (2010). SeDiCI (Servicio de Difusión de la Creación Intelectual): un recorrido de experiencias (2003-2011). *Jornada Virtual de Acceso Abierto*. Revisado Julio, 16, 2014 desde <http://hdl.handle.net/10915/27160>
8. DSpace 4.x Documentation - DSpace 4.x Documentation - DuraSpace Wiki. (n.d.). Revisado Julio, 16, 2014 desde <https://wiki.duraspace.org/display/DSDOC4x/DSpace+4.x+Documentation>
9. Overview (Java Platform SE 7). Revisado Julio, 16, 2014 desde <http://docs.oracle.com/javase/7/docs/api/>

10. Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1995). Design patterns: elements of reusable object-oriented software.
11. Extensions and Addons Work - DSpace - DuraSpace Wiki. (n.d.). Revisado Julio, 16, 2014 desde <https://wiki.duraspace.org/display/DSPACE/Extensions+and+Addons+Work>
12. Curation Task Cookbook - DSpace - DuraSpace Wiki. (n.d.). Revisado Julio, 16, 2014 desde <https://wiki.duraspace.org/display/DSPACE/Curation+Task+Cookbook>
13. Unified Modeling Language (UML). (n.d.). Revisado Julio, 16, 2014 desde <http://www.uml.org/#UML2.0>
14. Riehle, D. (1997). Cap 3.4. "Builder". A role-based design pattern catalog of atomic and composite patterns structured by pattern purpose (p. 14). Ubilab Technical Report 97.1. 1. Zürich, Switzerland: Union Bank of Switzerland.
15. Curation System - DSpace 4.x Documentation - DuraSpace Wiki. (n.d.). Revisado Julio, 16, 2014 desde <https://wiki.duraspace.org/display/DSDOC4x/Curation+System>
16. Reference Model for an Open Archival Information System (OAIS), CCSDS Magenta Book. Issue 1. . (n.d.). Revisado Julio, 16, 2014 desde <http://public.ccsds.org/publications/archive/650x0m2.pdf>
17. DCMI Metadata Terms. (n.d.). Revisado Julio, 16, 2014 desde <http://dublincore.org/documents/dcmi-terms/>
18. Chapter 2.1. DSpace System Documentation: Functional Overview. (n.d.). Revisado Julio, 16, 2014 desde http://dspace.org/sites/dspace.org/files/archive/1_5_2Documentation/ch02.html#N100A1
19. Validating CheckSums of Bitstreams - DSpace 4.x Documentation - DuraSpace Wiki. (n.d.). Revisado Julio, 16, 2014 desde <https://wiki.duraspace.org/display/DSDOC4x/Validating+CheckSums+of+Bitstreams#ValidatingCheckSumsofBitstreams-AutomatedChecksumCheckers'Results>
20. Rivest, R. (n.d.). The MD5 Message-Digest Algorithm. Revisado Julio, 16, 2014 desde <http://tools.ietf.org/html/rfc1321>
21. Sun, S., Reilly, S., Lannom, L., & Petrone, J. (n.d.). RFC 3652-Handle System Protocol (v2.1)- November 2003. Revisado Julio, 16, 2014 desde <http://www.ietf.org/rfc/rfc3652.txt>
22. De Giusti, M. R., Lira, A. J., Villarreal, G. L., Terruzzi, F. A., & Adorno, F. G. (2014, March 27). Preservación digital: un experimento con SEDICI-DSpace. *XX Asamblea General de ISTE (Puebla, México, 2014)*.



PRODUÇÃO CIENTÍFICA SOBRE PRESERVAÇÃO DIGITAL: LEVANTAMENTO BIBLIOGRÁFICO E DOCUMENTAL EM BASES DE DADOS¹

Danilo Formenton¹, Luciana de Souza Gracioso²

¹Mestrando em Ciência, Tecnologia e Sociedade pela Universidade Federal de São Carlos e Graduado em Biblioteconomia e Ciência da Informação pela Universidade Federal de São Carlos e-mail: formenton.danilo@gmail.com

²Doutora em Ciência da informação pela Universidade Federal Fluminense e Instituto Brasileiro de Informação em Ciência e Tecnologia e Docente do Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade da Universidade Federal de São Carlos e-mail: lgracioso@yahoo.com.br

RESUMO

Frente a crescente necessidade de criação e sistematização de práticas relacionadas à preservação digital este trabalho procura identificar, descrever e analisar um conjunto de registros obtidos sobre preservação digital em bases de dados com abrangência nacional e internacional, sendo elas: *ScienceDirect*, *Scopus*, *Web of Science*, *LISA* e *SciELO*. Objetiva indicar os períodos e as fontes em que se possa obter o maior número de registros acerca do tema preservação digital no intuito de promover um arranjo quantitativo sobre o assunto. Configura-se como uma pesquisa exploratória, de natureza bibliográfica e documental. Constitui o corpus de análise registros obtidos no período entre 2003 a 2013 para a consecução de um cenário mais atual referente ao tema. Verifica que a base *Web of Science* apresentou, no panorama internacional, o maior número de registros. No cenário internacional, no contexto do campo das Ciências Sociais, o período de 2009 a 2011 se destacou na produção de estudos, enquanto que as revistas científicas e as revisões se sobressaíram como principais fontes de informação. Conclui-se, através do pequeno volume de resultados constatado em certas bases, que torna-se indispensável o desenvolvimento de mais iniciativas nacionais neste domínio. Os resultados obtidos neste trabalho permitirá apoiar as comunidades interessadas e as áreas que compõem a Ciências Sociais como, a Ciência da Informação, na elaboração de pesquisas e práticas no âmbito da preservação digital, além de propiciar a exposição das possíveis tendências existentes, nos últimos anos, sobre a produção científica neste domínio, colaborando para o planejamento e atualização das organizações.

Palavras-chave: Preservação digital. Levantamento bibliográfico e documental. Base de dados. Ciência da Informação.

¹ Reprodução parcial do conteúdo e dos resultados presentes no trabalho sob o título de “A produção científica sobre preservação digital: levantamento bibliográfico e documental em bases de dados e repositórios institucionais” e que está atualmente submetido para o processo de avaliação pelos pares na revista Ponto de Acesso do Instituto de Ciência da informação-ICI, da Universidade Federal da Bahia (UFBA).



SCIENTIFIC PRODUCTION ABOUT DIGITAL PRESERVATION: BIBLIOGRAPHIC DOCUMENTARY SURVEY IN DATABASES

ABSTRACT

Regarding the growing need about creation and systematization of digital preservation practices, this paper aims to identify, describe and analyze a set of records about digital preservation obtained in national and international databases as: ScienceDirect, Scopus, Web of Science, LISA and SciELO. The analyses aimed to identify the sources and the periods in which it is possible to get the highest number of registrations about digital preservation topic in order to promote a quantitative arrangement on this subject. This work is classified as bibliographic and documentary exploratory research. The analysis sampling is constituted of records obtained between 2003 and 2013 in order to have a current scenario on the topic. It was verified that Web of Science database, presented the largest number of records in international arena. On the international scenario, in Social Sciences context, the period 2009-2011 was highlighted in the production of studies, review and scientific journals were the major source of information. In conclusion, because of the small volume of results recovered in certain databases, it becomes essential developing more national initiatives. The results of this work will assist interested communities and areas that integrate the Social Sciences, as Information Science, to develop research and practice in digital preservation field, and provide exposure of trends in recent years about scientific production, working to update and planning organizations.

Keywords: Digital Preservation. Bibliographic and documentary research. Databases. Information Science.

INTRODUÇÃO

A preservação digital, nos últimos anos, vem sendo alvo de investigação pela comunidade científica nacional e internacional conforme podemos constatar em Hedstrom (1998), Saramago (2002) e Arellano (2004). No contexto da Ciência da Informação, campo que trata da informação e do conhecimento, os estudos publicados em periódicos nacionais geralmente se pautam numa discussão mais conceitual sobre o tema, relatando o processo de preservação digital como o caminho mais adequado para a garantia da preservação e do acesso utilizável, à longo prazo, ao conteúdo das informações digitais. Russel e Sergeant (1999) consideram a preservação digital como o armazenamento, a manutenção e o acesso dos objetos digitais por longo prazo, usualmente como consequência da aplicação de uma ou mais estratégias de preservação digital, incluindo a preservação de tecnologia, a emulação tecnológica ou a migração de dados.

Entretanto, no panorama das publicações nacionais do campo da Ciência da Informação, são poucos os estudos que se dedicam inteiramente ao desenvolvimento de um levantamento bibliográfico e documental em bases de dados, com o intuito de identificar os períodos e tipos de fontes como forma de auxiliar o conjunto de diretrizes e políticas adotadas pelas organizações interessadas ou envolvidas com tal temática.

Além disso, baseado em estudos recentes produzidos neste campo, pontua-se que comparado a alguns países europeus como Portugal, que são referências em pesquisas e projetos na área de preservação, ainda é singela a contribuição brasileira no que diz respeito tanto a própria realização de estudos científicos sobre a preservação digital como a criação de iniciativas em prol da preservação e do acesso utilizável dos recursos de informação em meio digital por longo período.

Neste cenário, realizou-se a análise dos resultados advindos do levantamento em bases de dados com abrangência nacional e internacional em estudos científicos e que disponibilizam diferentes recursos para a elaboração de estratégias de busca ou refinamento de resultados, sendo elas: *ScienceDirect*, *Scopus*, *Web of Science* e *Library and Information Science Abstracts (LISA)*, para a obtenção de um panorama internacional sobre o assunto; e *Scientific Electronic Library Online (SciELO)*, para contextualizar a produção científica nacional referente ao tema.

O presente trabalho possui o propósito de identificar os períodos e os tipos de fontes onde se possa obter o maior número de estudos e informações acerca dos diversos aspectos envolvidos na preservação digital, a fim de poder servir como apoio em pesquisas e projetos nacionais no contexto das áreas que compreendem o campo das Ciências Sociais como, por exemplo, a Ciência da Informação. Os resultados ainda reproduzem as possíveis tendências e caminhos traçados, nos últimos anos, da produção científica neste domínio, contribuindo para o planejamento e atualização das organizações. Sendo assim, configura-se como pesquisa exploratória, onde o método de coleta de dados baseou-se em pesquisa bibliográfica e documental. Se utilizou de tabulação de dados e análise de conteúdo para a discussão dos resultados.

PROCEDIMENTOS METODOLÓGICOS

A partir de uma pesquisa de caráter livre, sem qualquer recorte temporal ou geográfico, sobre publicações quanto à preservação digital, foi possível construir um referencial teórico onde mediante uma revisão identificou-se as palavras-chave e os conceitos mais recorrentes além dos países que são referências na discussão e no desenvolvimento de iniciativas na área da preservação como é o caso, por exemplo, de Portugal. Desta forma, através da análise qualitativa do conteúdo dos resumos e das palavras-chave adotadas em publicações de periódicos nacionais do campo da Ciência da Informação, reconhecidos pelo Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), e também no periódico português *Cadernos de Biblioteconomia, Arquivística e Documentação (Cadernos BAD)*, importante veículo de divulgação científica e acesso à publicações nas áreas que abrange, foram criadas e elencadas um conjunto de trinta e seis expressões de busca nos idiomas português e inglês com ou sem operadores de busca *booleanos* de intersecção e união, possuindo assim diferentes níveis de especificidade.

Assim, se desenvolveu um levantamento bibliográfico e documental entre 2003 a 2013 (dez anos) sobre a preservação digital nas bases citadas anteriormente, com o objetivo de obter um cenário mais atual dos registros sobre esta temática de investigação. Os registros, após terem sido recuperados, foram organizados e inseridos em um conjunto de quadros que, em seguida, ao serem analisados ocorreu a coleta dos dados considerados pertinentes para a elaboração das tabelas expostas neste trabalho. Para o tratamento e a análise dos resultados com o propósito de criar as tabelas, adotaram-se os recursos disponíveis no software *Microsoft Office Excel* versão 2010, sendo que para o arredondamento de um número para um número especificado de dígitos (um, dois ou três, por exemplo) aplicou-se a função *ARRED*.

Durante o processo, a escolha do emprego ou não simultâneo de expressões nos idiomas português e inglês com ou sem o uso de operadores *booleanos* assim como dos recursos e filtros de busca disponíveis em cada base, se baseou nas opções onde se obteve o maior número de resultados. Embora durante todo o processo tenha ocorrido à indicação do período a ser abrangido para o levantamento, identificaram-se em certos casos registros recuperados não inclusos neste período. O processo foi realizado entre junho e julho do ano de 2013. Desta maneira, foram obtidos e expostos nas tabelas apenas dados parciais quanto aos registros referentes ao ano de 2013. Ademais, o acesso as bases se deu por meio do Portal de Periódicos da CAPES, com a inserção de suas respectivas denominações no campo *Buscar base*.

RESULTADOS RECUPERADOS NAS BASES DE DADOS COM ABRANGÊNCIA INTERNACIONAL

ScienceDirect¹

No processo de levantamento realizado nesta base de dados, obteve-se um conjunto elevado de resultados (888 registros), sendo que deste volume 75% correspondeu à categoria de assunto *All Sciences* (666 registros) e 25% a categoria *Social Sciences* (222 registros). Desta forma, a base *ScienceDirect* se constituiu no presente processo como a fonte onde se recuperou o quarto maior número de estudos produzidos no mundo sobre a preservação digital. Assim, pode-se estabelecer na Tabela 1, a partir do número total de resultados obtidos nesta base para estas categorias, a distribuição ao longo de onze anos da quantidade e subsequente frequência da ocorrência dos registros sobre os assuntos de preservação digital.

¹ Adotaram-se o recurso *Search* e posteriormente *All Sources*, sendo que todas as expressões de busca foram inseridas no campo *Search* na opção *Abstract, Title, Keywords* em *All Sources*. Para cada expressão, na opção *Subject*, empregou-se a categoria de assunto *All Sciences (Life Sciences, Health Sciences, Physical Sciences and Engineering e Social Sciences & Humanities)* na primeira busca e, depois, somente a categoria *Social Sciences* na segunda busca. Em ambas as buscas realizadas, se utilizaram os itens *Journals* e *All Books* na opção de *Include*. Além disso, aplicou-se *All sources* na opção *Source* e o valor *2003 to Present* em *Date Range*.

Tabela 1 - Distribuição da quantidade e frequência da ocorrência de registros obtidos para as categorias de assunto no período entre 2002-2013, base *ScienceDirect*.

Ano	Número total de registros na categoria de assunto <i>All Sciences (AS)</i> e frequência em porcentagem	Número total de registros na categoria de assunto <i>Social Sciences (SS)</i> e frequência em porcentagem	Número total de registros nas categorias de assunto <i>All Sciences (AS)</i> e <i>Social Sciences (SS)</i> e frequência em porcentagem
2002	3 (0,5%)	3 (1,4%)	6 (0,7%)
2003	41 (6,2%)	22 (9,9%)	63 (7,1%)
2004	52 (7,8%)	20 (9%)	72 (8,1%)
2005	34 (5,1%)	2 (0,9%)	36 (4,1%)
2006	53 (8%)	17 (7,7%)	70 (7,9%)
2007	71 (10,7%)	20 (9%)	91 (10,2%)
2008	77 (11,6%)	22 (9,9%)	99 (11,1%)
2009	67 (10,1%)	30 (13,5%)	97 (10,9%)
2010	49 (7,4%)	13 (5,9%)	62 (7%)
2011	70 (10,5%)	18 (8,1%)	88 (9,9%)
2012	79 (11,9%)	31 (14%)	110 (12,4%)
2013²	70 (10,5%)	24 (10,8%)	94 (10,6%)
TOTAL	666 (75% do número total de registros para ambas as categorias de assunto)	222 (25% do número total de registros para ambas as categorias de assunto)	888 (100%)

Fonte: Elaborado pelos autores.

Através da tabela foi possível observar que no número de registros obtidos somente para a categoria de assunto *Social Sciences*, o período entre os anos de 2011 a 2013 (73 registros, 32,9%) demonstrou um elevado índice de produtividade de estudos durante estes onze anos. Ademais, os anos de 2012, 2009 e 2013 (85 registros, 38,3%) se destacaram com as maiores quantidades e subsequentes frequências de ocorrência de registros. Apesar da existência de dados parciais referentes ao ano de 2013, torna-se notável o crescimento progressivo de iniciativas nos últimos anos, em especial a partir do ano de 2009, por parte do campo das Ciências Sociais, evidenciando assim que a preservação digital e os desafios que se propõe a amenizar ou resolver, vem sendo investigados pelas áreas do conhecimento que compreendem este campo científico.

Além disto, a Tabela 2 presente abaixo expressa o número total de resultados obtidos para os tipos de documento *Journal*, *Book* e *Reference Work* nas categorias de assunto *All Sciences* e *Social Sciences*, com indicação da quantidade e frequência da ocorrência de registros.

² Foram recuperados e expostos apenas dados parciais quanto aos registros referentes ao ano de 2013. Portanto, tal marcação se seguirá ao longo das demais tabelas, sendo que em todas as ocorrências de análises referentes à 2013 deverão ser levadas em conta este resultado parcial.

Tabela 2 - Distribuição da quantidade e frequência da ocorrência de registros obtidos para tipos de documento nas categorias de assunto, base *ScienceDirect*.

Número total de registros para <i>Journal</i> nas categorias de assunto <i>All Sciences (AS)</i> e <i>Social Sciences (SS)</i> , além de frequência em porcentagem		Número total de registros para <i>Book</i> nas categorias de assunto <i>All Sciences (AS)</i> e <i>Social Sciences (SS)</i> , além de frequência em porcentagem		Número total de registros para <i>Reference Work</i> nas categorias de assunto <i>All Sciences (AS)</i> e <i>Social Sciences (SS)</i> , além de frequência em porcentagem	
AS	SS	AS	SS	AS	SS
586 (74,6%)	199 (25,4%)	90 (81,8%)	20 (18,2%)	16 (53,3%)	14 (46,7%)
785 (AS e SS, 100%)		110 (AS e SS, 100%)		30 (AS e SS, 100%)	

Fonte: Elaborado pelos autores.

Nesta tabela se constatou que do número total de registros obtidos para *Journal* (785 registros), *Book* (110 registros) e *Reference Work* (30 registros), respectivamente 25,4% (199 registros), 18,2% (20 registros) e 46,7% (14 registros) corresponderam à categoria *Social Sciences*. Deste modo, as áreas que integram o campo da Ciências Sociais como, a área da Ciência da Informação, através deste levantamento podem assumir as revistas científicas, os livros e as obras de referência indexadas na base *ScienceDirect*, nesta ordem, como grandes fontes de informação para o desenvolvimento de pesquisas e práticas no âmbito da preservação digital.

Scopus³

Com relação ao levantamento efetuado nesta base de dados, obteve-se um conjunto grande de resultados (2683 registros), sendo que deste volume 64,3% equivaleu à categoria de assunto *All Sciences* (1726 registros) e 35,7% a categoria *Social Sciences e Humanities* (957 registros). Assim, a base *Scopus* se estabeleceu no presente processo de levantamento, como a fonte onde se obteve o terceiro maior número de estudos produzidos no mundo sobre a preservação digital. À vista disto, pode-se estabelecer na Tabela 3, por meio do número total de resultados obtidos nesta base para essas categorias, a distribuição ao longo de dez anos da quantidade e subseqüente frequência da ocorrência dos registros associados aos assuntos de preservação digital.

³ Adotaram-se o recurso *Search* e posteriormente *Document search*, sendo que todas as expressões de busca foram inseridas no campo *Search for* na opção *Article Title, Abstract, Keywords*. Para cada expressão na opção *Subject Areas*, selecionaram-se todas as áreas (*Life Sciences, Health Sciences,*

Tabela 3 - Distribuição da quantidade e frequência da ocorrência de registros obtidos para as categorias de assunto no período entre 2003-2013, base *Scopus*.

Ano	Número total de registros na categoria de assunto <i>All Sciences (AS)</i> e frequência em porcentagem	Número total de registros na categoria de assunto <i>Social Sciences & Humanities (SS&H)</i> e frequência em porcentagem	Número total de registros nas categorias de assunto <i>All Sciences (AS)</i> e <i>Social Sciences & Humanities (SS&H)</i> , além da frequência em porcentagem
2003	28 (1,6%)	16 (1,7%)	44 (1,6%)
2004	132 (7,6%)	42 (4,4%)	174 (6,5%)
2005	115 (6,7%)	51 (5,3%)	166 (6,2%)
2006	191 (11,1%)	117 (12,2%)	308 (11,5%)
2007	141 (8,2%)	99 (10,3%)	240 (8,9%)
2008	168 (9,7%)	87 (9,1%)	255 (9,5%)
2009	221 (12,8%)	124 (13%)	345 (12,9%)
2010	253 (14,7%)	131 (13,7%)	384 (14,3%)
2011	198 (11,5%)	115 (12%)	313 (11,7%)
2012	244 (14,1%)	158 (16,5%)	402 (15%)
2013 ⁴	35 (2%)	17 (1,8%)	52 (1,9%)
TOTAL	1726 (64,3% do número total de registros para ambas as categorias de assunto)	957 (35,7% do número total de registros para ambas as categorias de assunto)	2683 (100%)

Fonte: Elaborado pelos autores.

A partir da tabela se verificou que no número total de registros obtidos apenas para a categoria de assunto *Social Sciences & Humanities*, o período entre os anos de 2009 a 2011 (370 registros, 38,7%) possuiu um alto índice de produtividade de trabalhos durante estes dez anos. Os anos de 2012, 2010 e 2009 (413 registros, 43,2%) se destacaram com as maiores quantidades e frequências de ocorrência de registros. Desta forma, os resultados advindos da base *Scopus* reproduzem uma tendência similar a já constatada na base *ScienceDirect*, sobretudo, no que se refere aos anos de 2009 e 2012 onde estes se sobressaíram na produção de trabalhos ligados a preservação digital.

Além disso, pode-se constatar na Tabela 4, através do número total de resultados obtidos nesta base para as categorias de assunto *All Sciences* e *Social Sciences & Humanities*, a distribuição da quantidade e frequência da ocorrência de registros referentes aos assuntos de preservação digital nos diversos tipos de documentos.

⁴ *Physical Sciences* e *Social Sciences & Humanities*) na primeira busca e, depois, somente a área *Social Sciences & Humanities* na segunda busca. Em ambas as buscas efetuadas, escolheu-se a opção *All* em Document Type. Além disto, aplicou-se o valor *Published 2003 to Present* em *Date Range* no campo de *Limit to*

Tabela 4 - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os diversos tipos de documento nas categorias de assunto, base *Scopus*.

Número total de registros para os diversos tipos de documento na categoria de assunto <i>All Sciences (AS)</i> e frequência em porcentagem		Número total de registros para os diversos tipos de documento na categoria de assunto <i>Social Sciences & Humanities (SS&H)</i> e frequência em porcentagem		Número total de registros para os diversos tipos de documento nas categorias de assunto <i>All Sciences (AS)</i> e <i>Social Sciences & Humanities (SS&H)</i> , além de frequência em porcentagem	
<i>Article</i>	661 (38,3%)	<i>Article</i>	558 (58,31%)	<i>Article</i>	1219 (45,43%)
<i>Article in Press</i>	3 (0,17%)	<i>Article in Press</i>	1 (0,1%)	<i>Article in Press</i>	4 (0,15%)
<i>Conference Paper</i>	880 (50,98%)	<i>Conference Paper</i>	269 (28,11%)	<i>Conference Paper</i>	1149 (42,83%)
<i>Conference Review</i>	44 (2,55%)	<i>Conference Review</i>	19 (1,99%)	<i>Conference Review</i>	63 (2,35%)
<i>Editorial</i>	14 (0,81%)	<i>Editorial</i>	13 (1,36%)	<i>Editorial</i>	27 (1,01%)
<i>Note</i>	6 (0,35%)	<i>Note</i>	4 (0,42%)	<i>Note</i>	10 (0,37%)
<i>Review</i>	116 (6,72%)	<i>Review</i>	93 (9,72%)	<i>Review</i>	209 (7,79%)
<i>Short Survey</i>	2 (0,12%)	<i>Short Survey</i>	0 (0%)	<i>Short Survey</i>	2 (0,07%)
TOTAL	1726 (100%)	TOTAL	957 (100%)	TOTAL	2683 (100%)

Fonte: Elaborado pelos autores.

Na tabela pode ser observado que no número total de registros obtidos somente na categoria de assunto *Social Sciences & Humanities*, identificou-se altas quantidades e frequências de ocorrência de registros nos documentos *Article*, *Conference Paper* e *Review* (920 registros, 96,14%). Desta maneira, embora tratando-se do campo das Ciências Sociais e Humanas, novamente os artigos publicados em revistas científicas que abrangem as áreas que compõem estes campos, se sobressaem como principais fontes de informação na base *Scopus* para a obtenção de dados ou informações sobre a preservação digital.

Pode-se também averiguar na Tabela 5, a partir do número total de resultados obtidos nesta base para as categorias de assunto *All Sciences* e *Social Sciences & Humanities*, a distribuição da quantidade e frequência da ocorrência de registros relacionados aos assuntos de preservação digital no conjunto dos cinco primeiros países com maior volume de registros além do Brasil.

Tabela 5 - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os cinco primeiros países com maior volume de registros além do Brasil nas categorias de assunto, base *Scopus*.

País	Número total de registros na categoria de assunto <i>All Sciences (AS)</i> e frequência em porcentagem	Número total de registros na categoria de assunto <i>Social Sciences & Humanities (SS&H)</i> e frequência em porcentagem	Número total de registros nas categorias de assunto <i>All Sciences (AS)</i> e <i>Social Sciences & Humanities (SS&H)</i> , além de frequência em porcentagem
<i>Australia</i>	7 (0,53%)	5 (0,69%)	12 (0,59%)
<i>Austria</i>	131 (9,98%)	24 (3,33%)	155 (7,62%)
<i>Belgium</i>	3 (0,23%)	4 (0,56%)	7 (0,34%)
<i>Botswana</i>	1 (0,08%)	1 (0,14%)	2 (0,1%)
Brazil	14 (1,07%)	9 (1,25%)	23 (1,13%)
Canada	5 (0,38%)	35 (4,86%)	40 (1,97%)
<i>China</i>	4 (0,3%)	3 (0,42%)	7 (0,34%)
<i>France</i>	1 (0,08%)	0 (0%)	1 (0,05%)
<i>Finland</i>	0 (0%)	1 (0,14%)	1 (0,05%)
<i>Germany</i>	113 (8,61%)	29 (4,03%)	142 (6,98%)
<i>India</i>	0 (0%)	7 (0,97%)	7 (0,34%)
<i>Iran</i>	2 (0,15%)	2 (0,28%)	4 (0,2%)
<i>Italy</i>	2 (0,15%)	1 (0,14%)	3 (0,15%)
<i>Mexico</i>	4 (0,3%)	7 (0,97%)	11 (0,54%)
<i>Netherlands</i>	70 (5,33%)	28 (3,89%)	98 (4,82%)
<i>Nigeria</i>	0 (0%)	2 (0,28%)	2 (0,1%)
Portugal	56 (4,27%)	11 (1,53%)	67 (3,3%)
<i>South Africa</i>	1 (0,08%)	1 (0,14%)	2 (0,1%)
<i>Spain</i>	1 (0,08%)	1 (0,14%)	2 (0,1%)
<i>Sudan</i>	1 (0,08%)	1 (0,14%)	2 (0,1%)
<i>Sweden</i>	0 (0%)	3 (0,42%)	3 (0,15%)
<i>Switzerland</i>	3 (0,23%)	0 (0%)	3 (0,15%)
United Kingdom	211 (16,07%)	140 (19,44%)	351 (17,27%)
United States	683 (52,02%)	405 (56,25%)	1088 (53,52%)
TOTAL	1313 (64,58% do número total de registros para ambas as categorias de assunto)	720 (35,42% do número total de registros para ambas as categorias de assunto)	2033 (100%)

Fonte: Elaborado pelos autores.

Nesta tabela no que concerne ao número total de registros recuperados apenas na categoria *Social Sciences & Humanities*, os países *United States* (405 registros, 56,25%), *United Kingdom* (140 registros, 19,44%) e *Canada* (35 registros, 4,86%) demonstraram as maiores quantidades e subseqüentes frequências de ocorrência de registros. Ademais, *Brazil* expôs poucos registros (9 registros, 1,25%), se comparado ao volume obtido, por exemplo, para *Portugal* (11 registros, 1,53%). Desta forma, as comunidades envolvidas com o problema e que contribuem consideravelmente no desenvolvimento de estudos científicos e de práticas em prol da preservação e do acesso utilizável, por longo prazo, ao conteúdo dos recursos digitais, estão localizadas principalmente na Europa e na América do Norte e, em menor quantidade na América do Sul, onde o tema ainda é recente e carente de discussões teóricas e práticas.

Web of Science⁵

Em referência ao levantamento executado nesta base de dados, obteve-se um conjunto expressivo de resultados (1141543 registros), sendo que deste volume 93,83% correspondeu à categoria *All Citation Databases* (1071155 registros) e 6,17% a categoria *Social Sciences Citation Index* (70388 registros). Assim sendo, a base *Web of Science* se configurou no presente processo de levantamento, como a fonte onde se recuperou o maior número de estudos produzidos no mundo sobre o tema. Logo, pode-se estabelecer na Tabela 6, a partir do número total de resultados obtidos nesta base para estas categorias, a distribuição ao longo de dez anos da quantidade e subseqüente frequência da ocorrência dos registros relacionados aos assuntos de preservação digital.

⁵ Adotou-se o recurso *Search*, sendo que todas as expressões de busca foram inseridas no campo *Search* na opção *Topic*. Para cada expressão na opção de *Citation Databases*, selecionaram todas as bases (*Science Citation Index Expanded*, *Social Sciences Citation Index*, *Arts & Humanities Citation Index*, *Conference Proceedings Citation Index - Science* e *Conference Proceedings Citation Index - Social Science & Humanities*) na primeira busca e, depois, apenas a base *Social Sciences Citation Index* na segunda busca. Em ambas as buscas desenvolvidas, escolheu-se no campo de *Limits* o valor *From 2003 to 2013* em *Timespan*.

Tabela 6 - Distribuição da quantidade e frequência da ocorrência de registros obtidos para as categorias de assunto no período entre 2003-2013, base *Web of Science*.

Ano	Número total de registros na categoria <i>All Citation Databases (ACD)</i> e frequência em porcentagem	Número total de registros na categoria <i>Social Sciences Citation Index (SSCI)</i> e frequência em porcentagem	Número total de registros nas categorias <i>All Citation Databases (ACD)</i> e <i>Social Sciences Citation Index (SSCI)</i> , além da frequência em porcentagem
2003	76517 (7,143%)	3800 (5,399%)	80317 (7,036%)
2004	81936 (7,649%)	3686 (5,237%)	85622 (7,501%)
2005	87748 (8,192%)	4750 (6,748%)	92498 (8,103%)
2006	95559 (8,921%)	4837 (6,872%)	100396 (8,795%)
2007	102848 (9,602%)	5819 (8,267%)	108667 (9,519%)
2008	109221 (10,197%)	6738 (9,573%)	115959 (10,158%)
2009	115624 (10,794%)	7679 (10,91%)	123303 (10,801%)
2010	112498 (10,502%)	8907 (12,654%)	121405 (10,635%)
2011	122120 (11,401%)	9964 (14,156%)	132084 (11,571%)
2012	125471 (11,714%)	10454 (14,852%)	135925 (11,907%)
*2013	41613 (3,885%)	3754 (5,333%)	45367 (3,974%)
TOTAL	1071155 (93,83% do número total de registros para ambas as categorias)	70388 (6,17% do número total de registros para ambas as categorias)	1141543 (100%)

Fonte: Elaborado pelos autores.

Por meio da tabela identificou-se que no número total de registros obtidos apenas para a categoria de assunto *Social Sciences Citation Index*, o período entre os anos de 2009 a 2011 (26550 registros, 37,72%) exibiu um alto índice de produtividade de estudos durante estes dez anos. As maiores quantidades e frequências de ocorrência de registros se sucederam nos anos de 2012, 2011 e 2010 (29325 registros, 41,662%). Assim, os resultados que foram recuperados na base *Web of Science* seguem uma tendência semelhante a já observada na base *Scopus*, principalmente, em relação ao período entre 2009 a 2011 e aos anos de 2010 e 2012 onde estes se destacaram na produção de estudos vinculados a preservação digital.

Além disto, pode-se analisar na Tabela 7, através do número total de resultados obtidos nesta base para as categorias *All Citation Databases* e *Social Sciences Citation Index*, a distribuição da quantidade e frequência da ocorrência de registros relacionados aos assuntos de preservação digital nos diversos tipos de documentos.

Tabela 7 - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os diversos tipos de documento nas categorias de assunto, base *Web of Science*.

Número total de registros para os diversos tipos de documento na categoria All Citation Databases (ACD) e frequência em porcentagem		Número total de registros para os diversos tipos de documento na categoria Social Sciences Citation Index (SSCI) e frequência em porcentagem		Número total de registros para os diversos tipos de documento nas categorias All Citation Databases (ACD) e Social Sciences Citation Index (SSCI)	
Art Exhibit Review	8 (0,001 %)	Art Exhibit Review	0 (0%)	Art Exhibit Review	8 (0,001%)
Article	811344 (70,274%)	Article	55453 (76,029%)	Article	866797 (70,616%)
Biographical Item	72 (0,006%)	Biographical Item	16 (0,022%)	Biographical Item	88 (0,007%)
Bibliography	48 (0,004%)	Bibliography	20 (0,027%)	Bibliography	68 (0,006%)
Book Chapter	1355 (0,117%)	Book Chapter	52 (0,071%)	Book Chapter	1407 (0,115%)
Book Review	2124 (0,184%)	Book Review	1637 (2,244%)	Book Review	3761 (0,306%)
Correction	1531 (0,133%)	Correction	82 (0,112%)	Correction	1613 (0,131%)
Database Review	24 (0,002%)	Database Review	6 (0,008%)	Database Review	30 (0,002%)
Editorial Material	7615 (0,66%)	Editorial Material	1942 (2,663%)	Editorial Material	9557(0,779%)
Excerpt	4 (0%)	Excerpt	0 (0%)	Excerpt	4 (0%)
Fiction Creative Prose	4 (0%)	Fiction Creative Prose	0 (0%)	Fiction Creative Prose	4 (0%)
Film Review	4 (0%)	Film Review	0 (0%)	Film Review	4 (0%)
Hardware Review	4 (0%)	Hardware Review	0 (0%)	Hardware Review	4 (0%)
Letter	2622 (0,227%)	Letter	177 (0,243%)	Letter	2799 (0,228%)
Meeting Abstract	13105 (1,135%)	Meeting Abstract	1390 (1,906%)	Meeting Abstract	14495 (1,181%)
Music Performance Review	8 (0,001 %)	Music Performance Review	0 (0%)	Music Performance Review	8 (0,001%)
News Item	2042 (0,177%)	News Item	75 (0,103%)	News Item	2117 (0,172%)
Proceedings Paper	259192 (22,45%)	Proceedings Paper	2497 (3,424%)	Proceedings Paper	261689 (21,319%)
Record Review	52 (0,005%)	Record Review	0 (0%)	Record Review	52 (0,004%)
Reprint	152 (0,013%)	Reprint	24 (0,033%)	Reprint	176 (0,014%)
Review	53107 (4,6%)	Review	9538 (13,077%)	Review	62645 (5,104%)
Script	4 (0%)	Script	0 (0%)	Script	4 (0%)
Software Review	104 (0,009%)	Software Review	28 (0,038%)	Software Review	132 (0,011%)
Theater Review	8 (0,001 %)	Theater Review	0 (0%)	Theater Review	8 (0,001%)
TV Review Radio Review	8 (0,001 %)	TV Review Radio Review	0 (0%)	TV Review Radio Review	8 (0,001%)
TOTAL	1154541 (100%)	TOTAL	72937 (100%)	TOTAL	1227478 (100%)

Fonte: Elaborado pelos autores.

Nesta tabela constatou-se que em relação ao número total de registros obtidos somente na categoria de assunto *Social Sciences Citation Index*, os documentos *Article*, *Review* e *Proceedings Paper* (67488 registros, 92,53%), apresentaram as maiores quantidades e subsequentes frequências de ocorrência de registros. Em vista disto, seguindo a tendência presente nos levantamentos das bases *ScienceDirect* e *Scopus*, os artigos publicados em revistas científicas e as revisões, no campo das

Ciências Sociais, também se destacam na base *Web of Science* como grandes fontes de informação sobre a preservação digital.

Pode-se também desenvolver na Tabela 8, a partir do número total de resultados obtidos nesta base para as categorias *All Citation Databases* e *Social Sciences Citation Index*, a distribuição da quantidade e frequência da ocorrência de registros relacionados a preservação digital no conjunto dos cinco primeiros países com maior volume de registros além do Brasil.

Tabela 8 - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os cinco primeiros países com maior volume de registros além do Brasil nas categorias de assunto, base *Web of Science*.

País	Número total de registros na categoria <i>All Citation Databases</i> (ACD) e frequência em porcentagem	Número total de registros na categoria <i>Social Sciences Citation Index</i> (SSCI) e frequência em porcentagem	Número total de registros nas categorias <i>All Citation Databases</i> (ACD) e <i>Social Sciences Citation Index</i> (SSCI), além de frequência em porcentagem
<i>Argentina</i>	0 (0%)	1 (0%)	1 (0%)
<i>Australia</i>	16 (0%)	4418 (8,09%)	4434 (0,59%)
<i>Austria</i>	38 (0,01%)	0 (0%)	38 (0,01%)
<i>Belgium</i>	3 (0%)	1 (0%)	4 (0%)
Brazil	22303 (3,19%)	1039 (1,9%)	23342 (3,1%)
Canada	20 (0%)	4660 (8,53%)	4680 (0,62%)
<i>Cuba</i>	0 (0%)	1 (0%)	1 (0%)
England	59683 (8,54%)	8647 (15,83%)	68330 (9,07%)
<i>France</i>	14 (0%)	0 (0%)	14 (0%)
<i>Germany</i>	98750 (14,13%)	10 (0,02%)	98760 (13,11%)
<i>India</i>	7 (0%)	4 (0,01%)	11 (0%)
<i>Italy</i>	133 (0,02%)	3 (0,01%)	136 (0,02%)
<i>Japan</i>	82066 (11,74%)	0 (0%)	82066 (10,89%)
<i>Mexico</i>	1 (0%)	5 (0,01%)	6 (0%)
<i>Netherlands</i>	6 (0%)	12 (0,02%)	18 (0%)
<i>New Zealand</i>	0 (0%)	6 (0,01%)	6 (0%)
<i>Norway</i>	3 (0%)	0 (0%)	3 (0%)
<i>Peoples R China</i>	153300 (21,93%)	28 (0,05%)	153328 (20,35%)
Portugal	7743 (1,11%)	216 (0,4%)	7959 (1,06%)
<i>South Africa</i>	2 (0%)	5 (0,01%)	7 (0%)
<i>Spain</i>	5 (0%)	32 (0,06%)	37 (0%)
<i>Sudan</i>	1 (0%)	0 (0%)	1 (0%)
<i>Sweden</i>	2 (0%)	1 (0%)	3 (0%)
<i>Taiwan</i>	26 (0%)	2547 (4,66%)	2573 (0,34%)
USA	274819 (39,32%)	32977 (60,38%)	307796 (40,85%)

TOTAL	698941 (92,75% do número total de registros para ambas as categorias)	54613 (7,25% do número total de registros para ambas as categorias)	753554 (100%)
-------	---	---	---------------

Fonte: Elaborado pelos autores.

Foi possível verificar através da tabela que no número total de registros obtidos somente na categoria de assunto *Social Sciences Citation Index*, os países *USA* (32977 registros, 60,38%), *England* (8647 registros, 15,83%) e *Canada* (4660 registros, 8,53%) possuíram as mais elevadas quantidades e subsequentes frequências de ocorrência de registros. Ademais, *Brazil* expôs grandes quantidades e frequências de ocorrência de registros (1039 registros, 1,9%), se comparado ao volume de registros recuperados, por exemplo, para *Portugal* (216 registros, 0,4%). Apesar de não incluir o campo das Humanas e abranger apenas as Ciências Sociais, na base *Web of Science* similar ao ocorrido no levantamento da base *Scopus*, nas regiões da Europa e da América do Norte estão as comunidades que mais cooperam e contribuem em pesquisas sobre a preservação digital.

LISA⁶

No levantamento elaborado nesta base de dados, obteve-se um conjunto relativamente expressivo de resultados (52171 registros), sendo que deste volume 99,66% equivaleram a Periódicos acadêmicos (51993 registros), 0,33% a Procedimentos e Trabalhos de Conferência (170 registros) e, por fim, 0,02% a Outras Fontes (8 registros). Semelhante ao ocorrido nas bases *ScienceDirect*, *Scopus* e *Web of Science*, no cenário do campo das Ciências Sociais, as revistas científicas também se configuraram como principais fontes nesta base para a recuperação de informações sobre a preservação digital às comunidades interessadas. Deste modo, a base *LISA* se consolidou no presente processo de levantamento, como a fonte onde se recuperou o segundo maior número estudos produzidos no mundo quanto a preservação digital. Isto posto, pode-se estabelecer na Tabela 9, a partir do número total de resultados obtidos nesta base, a distribuição ao longo de dez anos da quantidade e subsequente frequência da ocorrência dos registros associados aos assuntos de preservação digital.

⁶ Adotaram-se o recurso de Busca avançada em Avançado, sendo que todas as expressões de busca foram inseridas na opção Assunto principal (todos) - SU. Para o campo Data de publicação escolheu-se a opção Intervalo de data específico e depois incluídos os valores entre os anos de 2003 e 2013. Não se utilizaram os filtros de Limitar a: Revisado por especialistas e Duplicatas: Incluir documentos duplicados. Ademais, foram selecionados todas as opções de fontes em Tipos de fontes, todas as opções de documentos em Tipos de documento e todos os idiomas disponíveis em Idiomas.

Tabela 9 - Distribuição da quantidade e frequência da ocorrência de registros obtidos no período entre 2003-2013, base *LISA*.

Ano	Número total de registros e frequência em porcentagem
2003	4400 (8,43%)
2004	4527 (8,68%)
2005	5790 (11,1%)
2006	6260 (12%)
2007	4397 (8,43%)
2008	4720 (9,05%)
2009	5994 (11,49%)
2010	5736 (10,99%)
2011	5446 (10,44%)
2012	4526 (8,67%)
*2013	384 (0,74%)
TOTAL	52180 (100%)

Fonte: Elaborado pelos autores.

Na tabela observou-se que do número total de resultados obtidos (52180 registros) o período entre os anos de 2009 a 2011 (17176 registros, 32,92%) exibiu um alto índice de produtividade de trabalhos durante estes dez anos, enquanto que os anos de 2006, 2009 e 2005 (18044 registros, 34,59%) demonstraram as maiores quantidades e frequências de ocorrência de registros. Os resultados obtidos na base *LISA* também refletiram a tendência presente nas bases *Scopus* e *Web of Science*, sobretudo, a grande produção de estudos sobre preservação digital no ano de 2009 e no período entre 2009 a 2011. Assim, através do conjunto de resultados advindos destas bases, as comunidades interessadas e as áreas que integram o campo da Ciência da Informação como, a Biblioteconomia e a Arquivologia, podem se pautar nestes anos, períodos e fontes de informação identificadas para o desenvolvimento de suas futuras pesquisas e práticas em prol da preservação e do acesso, à longo prazo, ao conteúdo dos recursos digitais.

RESULTADOS RECUPERADOS NAS BASES DE DADOS COM ABRANGÊNCIA NACIONAL

SciELO⁷

Em relação ao levantamento executado nesta base de dados, recuperou-se um número singelo de resultados (28 registros). Por conseguinte, pode-se estabelecer na Tabela 14, a partir do número total de resultados obtidos nesta base, um *ranking* dos periódicos nacionais de diversas áreas com maior número e frequência de ocorrência de registros para os assuntos sobre a preservação digital.

Tabela 14 - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os diversos periódicos nacionais de diferentes áreas, *SciELO*.

Posição no <i>ranking</i> e título do periódico	Número total de registros	Frequência em porcentagem do número total de registros
1. Ciência da Informação	11	39,286%
2. Revista Árvore	3	10,714%
3. Revista de Antropologia	2	7,143%
4. <i>Brazilian Journal of Oceanography</i>	2	7,143%
5. Ciência Rural	2	7,143%
6. Engenharia Agrícola	2	7,143%
7. Anais da Academia Brasileira de Ciências	1	3,571%
8. Fisioterapia em Movimento	1	3,571%
9. Revista Brasileira de Coloproctologia	1	3,571%
10. Revista Ceres	1	3,571%
11. Revista Brasileira de Engenharia Agrícola e Ambiental	1	3,571%
12. Revista de Odontologia da UNESP	1	3,571%
TOTAL	28	100%

Fonte: Elaborado pelos autores.

A partir da tabela observou-se que a maioria dos periódicos identificados com exceção do primeiro, possuiu uma quantidade e frequência de ocorrência de registros bem aproximada. Além disso, o periódico Ciência da Informação (11 registros, 39,286%) que abrange o campo da Ciência da Informação e da informação em geral bem como o periódico Revista Árvore (3 registros, 10,714%) que compreende o campo da Ciência Florestal e áreas afins, se destacaram no presente levantamento com as maiores quantidades e frequências de ocorrência de registros. Assim, percebe-se que a preservação digital é um tema de pesquisa multidisciplinar, pois não é alvo de investigação de somente as áreas que lidam com a gestão da informação e do conhecimento, mas também aquelas com origens nos campos das Ciências Biológicas e Exatas. Logo, as organizações envolvidas com esta temática podem então adotar como estratégia de pesquisa a não delimitação a áreas específicas.

⁷ Adotou-se a opção Pesquisa artigos, sendo que todas as expressões de busca foram inseridas no campo Entre com uma ou mais palavras. Para cada expressão utilizou-se o campo Método na opção Integrada e o campo Onde na opção Brasil. Com a recuperação de resultados, escolheu-se a opção Resumo no campo de Índices. Durante o processo não se identificou uma opção para a seleção do período a ser abrangido na recuperação dos estudos presentes na base de dados.

Ademais, pode-se demonstrar na Tabela 15, através do número total de resultados obtidos nesta base, a distribuição ao longo de dezesseis anos da quantidade e subsequente frequência da ocorrência de registros ligados aos assuntos de preservação digital.

Tabela 15 - Distribuição da quantidade e frequência da ocorrência de registros obtidos no período entre 1997-2013, *SciELO*.

Ano	Número total de registros	Frequência em porcentagem do número total de registros
1997	2	7,143%
1998	0	0%
1999	1	3,571%
2000	0	0%
2001	1	3,571%
2002	2	7,143%
2003	0	0%
2004	4	14,286%
2005	2	7,143%
2006	2	7,143%
2007	0	0%
2008	2	7,143%
2009	3	10,714%
2010	3	10,714%
2011	3	10,714%
2012	2	7,143%
*2013	1	3,571%
TOTAL	28	100%

Fonte: Elaborado pelos autores.

Nesta tabela também se constatou um volume aproximado de estudos produzidos ao longo destes dezesseis anos, com exceção do ano de 2004 (4 registros, 14,286%) onde ocorreu a maior quantidade e subsequente frequência de incidência de registros. Os períodos de maior produtividade observados foram entre os anos de 2009 a 2011 (9 registros, 32,142%) e entre 2003 a 2005 (6 registros, 21,429%). Desta maneira, o ano de 2004 se destaca na produção de artigos científicos juntamente com o período de 2009 a 2011, onde podemos verificar uma produção ainda singela, mas crescente e constante de estudos sobre a preservação digital.



CONSIDERAÇÕES FINAIS

A partir dos resultados obtidos no levantamento presente neste trabalho, no panorama científico internacional, a *Web of Science*, *LISA*, *Scopus* e *ScienceDirect*, configuraram-se, nesta ordem, como sendo as bases onde se recuperou os maiores números de estudos produzidos no mundo sobre a preservação digital. Nestas bases, no contexto do campo das Ciências Sociais, o período entre os anos de 2009 a 2011 e o ano de 2012 se destacaram na produção estudos sobre o tema, posto que as revistas científicas e as revisões que abrangem as áreas que compõem este campo, se sobressaíram como principais fontes para a obtenção de dados e informações. As comunidades presentes na Europa (Inglaterra) e na América do Norte (Estados Unidos e Canadá), envolvidas com problemática da preservação e do acesso à longo prazo ao conteúdo das informações digitais, são as que mais contribuem no desenvolvimento de estudos e iniciativas neste domínio.

Além disso, no cenário científico nacional sob o âmbito das publicações de diferentes áreas do conhecimento, o periódico *Ciência da Informação* e o ano de 2004 se consolidaram como grandes fontes de informação para a obtenção de estudos sobre a preservação digital. Apesar da grande quantidade de resultados recuperados nas diversas bases de dados levantadas, através da *SciELO* se verificou que o assunto ainda se caracteriza por ser novo e carente de iniciativas com relação ao contexto nacional.

Portanto, torna-se fundamental a realização de mais pesquisas sobre esta temática em benefício da comunidade científica nacional e internacional além da própria sociedade. A identificação deste conjunto de dados nestas bases de dados, poderá auxiliar as comunidades interessadas e as áreas que integram o campo da Ciências Sociais como, a *Ciência da Informação*, no desenvolvimento de pesquisas e práticas em prol da preservação e do acesso, por longo prazo, aos conteúdos digitais, além de apresentar as possíveis tendências existentes, nos últimos anos, quanto a produção científica neste domínio, colaborando para o planejamento e a respectiva atualização das organizações.

REFERÊNCIAS

ARELLANO, M. Á. M. Preservação de documentos digitais. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 15-27, maio/ago. 2004. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/305/1452>>. Acesso em: 10 jul. 2014.

CADERNOS DE BIBLIOTECONOMIA, ARQUIVÍSTICA e DOCUMENTAÇÃO (CADERNOS BAD). Disponível em: <http://www.apbad.pt/Edicoes/Edicoes_Cadernos.htm>. Acesso em: 10 jul. 2014.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). **Portal de Periódicos**. Disponível em: <<http://www.periodicos.capes.gov.br/>>. Acesso em: 10 jul. 2014.

HEDSTROM, M. Digital preservation: a time bomb for digital libraries. **Computer and**



the Humanities, v. 31, p. 189-202, 1998. Disponível em:
<<http://deepblue.lib.umich.edu/bitstream/handle/2027.42/42573/?sequence=1>>.
Acesso em: 10 jul. 2014.

LIBRARY AND INFORMATION SCIENCE ABSTRACTS (LISA). Disponível em:
<<http://search.proquest.ez31.periodicos.capes.gov.br/lisa/index>>. Acesso em: 10 jul.
2014.

RUSSEL, K; SERGEANT, D. The Cedars project: implementing a model for
distributed digital archives. **RLG DigiNews**, v. 3, n. 3, jun. 1999. Disponível em:
<<http://worldcat.org/arcviewer/2/OCC/2009/07/29/H1248898656915/viewer/file2.html>
>. Acesso em: 10 jul. 2014.

SARAMAGO, M. L. Preservação digital a longo prazo: boas práticas e estratégias.
Cadernos BAD, Lisboa, n. 2, p. 54-68, 2002. Disponível em:
<<http://www.apbad.pt/CadernosBAD/Caderno22002/Saramago.pdf>>. Acesso em: 10
jul. 2014.

SCIENCEDIRECT. Disponível em: <<http://www-sciencedirect-com.ez31.periodicos.capes.gov.br/>>. Acesso em: 10 jul. 2014.

SCIENTIFIC ELECTRONIC LIBRARY ONLINE (SCIELO). Disponível em:
<<http://www.scielo.org/php/index.php>>. Acesso em: 10 jul. 2014.

SCOPUS. Disponível em: <<http://www-scopus-com.ez31.periodicos.capes.gov.br/home.url>>. Acesso em: 10 jul. 2014.

WEBOFSCIENCE. Disponível em:
<http://apps.webofknowledge.com.ez31.periodicos.capes.gov.br/WOS_GeneralSearch_input.do?product=WOS&search_mode=GeneralSearch&SID=3DNI7L4GiboaLBM1jeA&preferencesSaved=>>. Acesso em: 10 jul. 2014.



CURADORIA DIGITAL DE DADOS DE PESQUISA: O CENÁRIO DAS FUNDAÇÕES DE AMPARO À PESQUISA NO BRASIL

*Prof. Dra. Laura Vilela Rodrigues Rezende*¹, *Prof. Dra. Sonia Cruz Riascos*²

¹Universidade Federal de Goiás, lauravil.rr@gmail.com,

²Universidade Federal de Pernambuco, sonia.cruzriascos@gmail.com.

RESUMO

A presente pesquisa trata da curadoria digital de dados de pesquisa no âmbito da comunicação científica brasileira, destacando a importância do uso, da preservação e da custódia em acesso aberto. Apresenta-se breve histórico da comunicação científica destacando a invasão de novos recursos eletrônicos e terminologias, com atenção especial ao acesso aberto (*Open Access*) e seus desdobramentos. Discorre-se sobre o cenário brasileiro de mecanismos de registro, disseminação e disponibilização da produção científica, enfatizando a questão da gestão dos dados digitais de pesquisa, em especial aquelas financiadas com recursos públicos. Sabe-se que, ao longo do tempo, a ciência teve como foco a busca por tecnologias que pudessem decifrar e dimensionar o mundo. Como consequência desta busca incessante, tem-se grande volume de dados gerados em laboratórios e institutos de pesquisa abastecendo planilhas infindáveis. Diante disto, pode-se afirmar que a ciência vive hoje um paradigma que vem alterando os rumos da pesquisa científica, a E-Science. Questiona-se, então: Como aproveitar de forma remota e multidisciplinar a quantidade cada vez maior de dados produzidos durante a realização das investigações? De fato, a consulta aos dados de pesquisa ora concluídas, por outras investigações em elaboração, ou ainda, em vias de serem realizadas depende de uma efetiva gestão desses dados para que possam ser armazenados em bases confiáveis garantindo a recuperação ao longo do tempo. Na tentativa de propor ações para gerir dados de pesquisa, tem-se a curadoria digital que se relaciona com o modo de disponibilizá-los em formato digital para acesso futuro. A ideia é tornar esses dados acessíveis, interpretáveis, reutilizáveis e confiáveis para pesquisas ainda não realizadas. Frente ao contexto exposto, o objetivo desta pesquisa foi o de realizar o mapeamento das informações disponíveis sobre investigações científicas financiadas pelas Fundações de Amparo à Pesquisa no Brasil, para identificar o tipo de informação disponível e, em especial, verificar se os dados que fizeram parte dos corpos das pesquisas se encontram disponibilizados para o público nos sites de tais Fundações. Tratou-se de pesquisa quantitativa, com abordagem descritiva, que utilizou-se da técnica do levantamento de dados, ao nível nacional, nos sites das Fundações Brasileiras de Apoio à Pesquisa. As Fundações de Apoio são instituições criadas com a finalidade de dar suporte aos projetos de pesquisa, ensino, extensão e de desenvolvimento institucional, científico e tecnológico, de interesse das Instituições



Federais de Ensino Superior (IFES), como também às Instituições de Pesquisa. São credenciadas ao Ministério da Educação e ao Ministério da Ciência e Tecnologia do Brasil. Ao final da investigação, os resultados obtidos se encontram apresentados com as considerações necessárias acerca do tema para que se possa aperfeiçoar o tratamento de dados digitais e a disponibilização de informações científicas, a exemplo da implementação de sistemas de gestão desses ativos, como os repositórios digitais, no cenário nacional. Tal medida visa favorecer a oferta de subsídios informacionais para o acesso, o uso e a reutilização para o avanço de pesquisas futuras. Ademais, a ideia de curadoria digital se estende além do controle do repositório que arquiva os recursos e envolve a atenção do criador do conteúdo e dos usuários futuros, indicando-se, portanto, a garantia da preservação e da custódia em ambiente digital, para uso e usufruto contínuo de pesquisadores e interessados.

Palavras-chave: Curadoria digital – Dados de pesquisa - Preservação digital - Acesso aberto

INTRODUÇÃO

A partir do surgimento dos meios de comunicação e da difusão do conhecimento científico, - que inicialmente foram alavancados pela invenção da prensa de caracteres móveis de Gutemberg, causando “grande impacto na divulgação de informações”, devido ao fato de que sua “capacidade de multiplicar os exemplares de um livro, - representou um passo importante rumo a uma disseminação melhor e mais rápida das pesquisas” (MEADOWS, 1999, p.4). Seguindo as palavras de Pinheiro e Kuramoto (2012, p. 307) pode-se afirmar que:

A comunicação foi invadida por novos recursos eletrônicos e terminologia não compreendida e definida. Sítios, portais, twitter, facebook e blogs representam e ampliam as possibilidades não somente de comunicar, mas de disseminar e disponibilizar informações na internet, por sua vez adotando técnicas inovadoras da arquitetura da informação. Novos são os mecanismos de registro, disseminação e disponibilização da produção científica – bibliotecas digitais e virtuais, repositórios institucionais e temáticos, portais de periódicos científicos eletrônicos – nos quais protocolos e compartilhamento de dados buscam facilitar as ações da ciência e tecnologia.



Segundo Mueller (2007, p. 137):

A comunicação informal e formal entre pesquisadores e entre esses e a sociedade como um todo foi muito afetada pela tecnologia da informação e comunicação. No Brasil, essa tecnologia começou a ser utilizada nas décadas de 1960 e 1970, mas só se tornou realmente comum nas atividades científicas das universidades a partir da década de 1990. Periódicos eletrônicos, correio eletrônico, acesso à internet, vídeo conferências e até mesmo *blogs e podcasts*, entre outras iniciativas, modificaram profundamente a comunicação científica, não apenas as publicações das revistas, mas também maneiras em que se realizam interações pessoais e o processo de produção das revistas científicas. A facilidade de acesso à internet modificou o papel das bibliotecas, ampliando as fontes disponíveis para consulta.

A crise dos periódicos, movimento caracterizado pelo alto custo dos periódicos científicos e sua inacessibilidade aos menos favorecidos, fez surgir um movimento em favor do acesso irrestrito aos produtos de pesquisas científicas através da internet, segundo afirma Costa (2005, p.175) “na verdade, a questão do preço das assinaturas tem sido uma das principais, se não a principal motivação para os movimentos do acesso aberto e dos arquivos abertos”.

Conforme Varela, Barbosa e Guimarães (2009, p. 129):

[...] a iniciativa dos Arquivos Abertos – *Open Archive Initiative (OAI)* [...], surge como uma tentativa na solução para o alto custo das assinaturas de periódicos científicos, a lentidão na disseminação da produção científica, as barreiras impostas pela legislação de direitos autorais; e a necessidade de reduzir o fosso científico, tecnológico e cultural que se instalou entre países ricos, emergentes e pobres, em função das disparidades educacionais e culturais, para as quais contribuem a centralidade, nas regiões mais prósperas, da produção e distribuição do conhecimento, reduzindo a possibilidade do sujeito das regiões periféricas de ampliar suas competências cognitivas e de aprendizagem, situação que contribui significativamente para acirrar o quadro de exclusão e desigualdade social.

A tendência principal da comunicação eletrônica hoje parece ser em direção a uma filosofia aberta, levantando-se questões sobre software aberto, acesso aberto, arquivos abertos, direito autoral, dentre outras.

Ainda em relação a temática do acesso aberto, recentemente a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), em sua publicação: Informe de la consulta latino-americano y del Caribe sobre Acceso Abierto a información e investigación científica: concepto e políticas, resultado da Congregação realizada em Kingston no ano de 2013, convocou diversas autoridades ao nível mundial para debater o referido tema, em virtude da necessidade de se chegar a consensos que permitam a implementação de ações e políticas de Acesso Aberto, em especial na América Latina e Caribe, localidades representadas nessa ocasião pela Congregação. (ORGANIZAÇÃO, 2013, tradução nossa).



Em relação aos benefícios proporcionados pelo movimento de acesso aberto, destaca-se sua função de dinamização e difusão do conhecimento, que apesar de apresentar barreiras físicas e econômicas, nota-se “a possibilidade de elevação da capacidade de produção de novos conhecimentos, os quais por sua vez, serão convertidos em benefícios para a sociedade como um todo.” (LEITE, 2012, p.414-415).

Sabe-se que no século XX, a maioria dos dados utilizados para a elaboração de teorias científicas hoje estão inacessíveis. As mudanças ocorridas nos formatos e mídias acabou por tornar ilegível um grande volume de dados criados no passado. Hoje, são muitos os desafios em torno da origem remota dos dados além do acesso comunitário a dados distribuídos. O presente estudo apresenta um mapeamento das informações sobre investigações científicas financiadas pelas Fundações de Amparo à Pesquisa no Brasil disponíveis em seus websites. O conceito que fundamenta esta investigação é o de curadoria digital, o qual se estende para além do controle do repositório que arquiva os dados de pesquisa envolvendo também a atenção do criador do conteúdo e dos usuários futuros, indicando, portanto, a garantia da preservação e da custódia em ambiente digital, para uso e usufruto contínuo de pesquisadores e interessados.

REFERENCIAL TEÓRICO

Dados de pesquisa: conceitos e reflexões atuais

Renato Cruz (2011, p.25) afirma que:

Uma visão antiga de ciência e tecnologia, o chamado modelo linear, dizia que, sem preocupação com aplicações práticas, a pesquisa básica gerava conhecimento e depois era transmitido para a pesquisa aplicada e para o desenvolvimento, que transformariam tal conhecimento em produto.

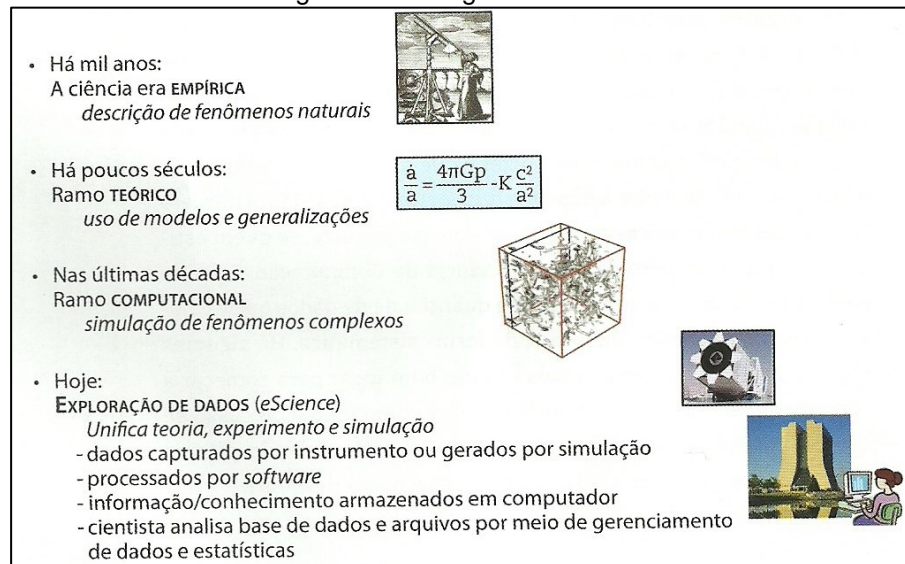
No entanto, ele acrescenta ainda que “há pesquisas que são básicas e, ao mesmo tempo, aplicadas”. Isto quer dizer que geram conhecimento teórico e aplicações práticas. Um aspecto importante no tocante à ciência nos dias atuais é que boa parte dela é alimentada pela tecnologia. Abbot (2011, p.133) afirma que a Tecnologia da Informação (TI) afetou a comunidade científica de duas maneiras. Primeiramente transformando o armazenamento genérico e a computação numa *commodity*. Esta parece ser uma opção razoável para as tarefas científicas que podem ser realizadas por meio de serviços genéricos e centralizados, como a computação em nuvens. A segunda maneira pela qual a TI afetou a comunidade científica, acrescenta o autor, foi por meio da personalização radical. Existe hoje uma rede de serviços de conhecimento distribuídos massivamente cujo desafio é controlá-la. Os cientistas podem criar suas próprias nuvens de computação devido ao acesso pessoal a terabytes de armazenamento e teraflops de computação. É possível visualizar neste cenário uma gama de redes informais de serviços potenciais e uma estreita parceria entre aqueles que criam conhecimento e os que o usam.

Gray (2007 *apud* HEY, TANSLEY, TOLLE, 2011, pág.18, tradução nossa) afirma que:

hoje, o novo modelo de se fazer Ciência prevê a captura de dados por instrumentos ou sua geração através de simulações, seguida por um processamento por software e armazenamento da informação ou conhecimento resultantes em computadores. Os cientistas só veem seus dados numa etapa bastante avançada dessa sequência.

Observando a figura a seguir, ficam evidentes as modificações sofridas pela ciência ao longo do tempo.

Figura 1: Paradigmas da Ciência



Fonte: (GRAY, 2007 *apud* HEY, TANSLEY, TOLLE, 2011, pág.18, tradução nossa)

Cunha (2010) discorre sobre a atual ciência eletrônica afirmando que essa área é, geralmente, composta por bases de dados numéricos e diferentes conjuntos de resultados das pesquisas realizadas nos institutos, faculdades e departamentos.

Junior (*apud* HEY, TANSLEY, TOLLE, 2011, pág.7, tradução nossa) discorre sobre a ciência afirmando que “uma de suas principais características está relacionada com medidas: para estudar um certo fenômeno, medidas objetivas são realizadas sobre amostras ligadas ao fenômeno através de instrumentos (sensores) especializados”. Ele acrescenta ainda que “a evolução da ciência está profundamente ligada à evolução dos instrumentos que permitem a realização de observações”. Ponderando suas considerações ele salienta que “o avanço recente da tecnologia levou a um deslocamento no gargalo para o avanço científico: em vez de a ciência não avançar devido à escassez de dados, hoje em dia ela frequentemente encontra dificuldades em avançar por seu excesso”. Finalizando, “[...] a pesquisa moderna compartilha uma característica comum: a capacidade de gerar e armazenar dados em uma escala sem precedentes e muito além da capacidade humana de análise”.



Ainda nessa publicação, para Junior:

Uma nova área científica nasceu nos últimos dez anos. Como ocorre frequentemente na história da ciência, ela nasceu em diferentes lugares, com diferentes nomes, mas da necessidade comum de se enfrentar o dilúvio de dados: “*e-Science*”, “ciência orientada por dados” (*data-driven science*), “computação fortemente orientada a dados” (*data-intensive computing*), “ciberinfraestrutura” (*cyberinfrastructure*), “quarto paradigma” (*fourth paradigm*), “dos dados ao conhecimento” (*from data to knowledge*). (HEY, TANSLEY, TOLLE, 2011, tradução nossa).

Curadoria Digital de dados de pesquisa

Na perspectiva de Sayão e Sales (2012), o principal objetivo do conceito de curadoria digital dos dados de pesquisa, configura-se na necessidade de preservar não somente os dados, mas também sua capacidade de transmitir conhecimentos para uso futuro por parte de seus interessados.

Gray (2007 *apud* HEY, TANSLEY, TOLLE, 2011, tradução nossa) afirma que “é preciso se fazer mais em termos de produzir ferramentas para o ciclo completo da pesquisa – captura dos dados e sua curadoria à análise e visualização dos mesmos”. Ele ainda salienta que:

Além da escassez observada em ferramentas para a captura, curadoria e análise de dados, a divulgação dos resultados de pesquisa parte do pressuposto de que todo projeto se encerra em algum momento e não está claro o que ocorre com os dados utilizados. Os dados coletados ao longo de todo o processo da pesquisa geralmente não são curados e nem publicados sistematicamente.

Uma consideração importante e que é consenso quando o assunto se trata de curadoria digital, relaciona-se ao fato de que o conceito é posterior à definição de preservação digital logo, uma vez que esta última se refere às atividades de gerenciamento necessárias para garantir acesso continuado aos materiais digitais pelo tempo que for preciso, pode-se afirmar que não existem ações de curadoria sem preservação digital.

Afirma-se, também, que curadoria digital está relacionada com o modo de tornar dados disponíveis em formato digital para acesso futuro. A ideia é fazer com que esses dados possam ser acessíveis, interpretáveis, reutilizáveis e confiáveis (NEUROTH; STRATHMANN; OBWALD; LUDWIG, 2013).

O *Digital Curation Center* (2014) afirma que a curadoria digital envolve manutenção, preservação e atribuição de valor a dados digitais de pesquisa durante todo o seu ciclo de vida. Ressalta ainda que o gerenciamento efetivo de dados de pesquisa durante todo o seu ciclo diminui os riscos de obsolescência e desvalorização da pesquisa ao longo do tempo.



Lee e Tibbo (2007) afirmam que a curadoria digital é responsável por:

- a) impulsionar a disponibilidade de dados para a comunidade científica;
- b) proporcionar armazenamento redundante;
- c) realizar transformações de dados; e para alguns tipos de materiais, o compromisso de preservação a longo prazo;
- d) disponibilizar dados digitais autênticos para serem reproduzidos e reutilizados;
- e) desenvolvimento de repositórios digitais confiáveis e duráveis;
- f) princípios de criação e captura de metadados de dados multimídia;
- g) uso de padrões abertos para formatos e conversão de arquivos.

Observa-se nas palavras de Corrêa e Bertocchi (2012, p.29) um relato de evolução e contextualização da prática da curadoria digital:

Na medida em que ocorre a expansão da sociedade digitalizada, o termo curadoria passa a ser utilizado para uma diversidade de ações que envolvem organização de dados a partir de critérios ou recortes. Nesse contexto, a “curadoria de informação” assume uma ideia muito mais de organização que de inauguração de uma nova proposta ou visão de mundo.

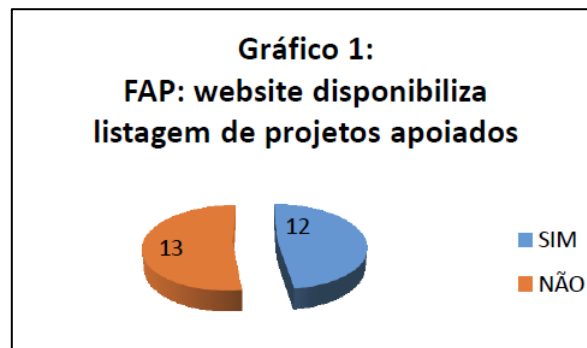
MATERIAIS E MÉTODOS

As Fundações de Amparo à Pesquisa (FAP) brasileiras são instituições criadas com a finalidade de dar suporte aos projetos de pesquisa, ensino, extensão e de desenvolvimento institucional, científico e tecnológico, de interesse das Instituições Federais de Ensino Superior (IFES), como também às Instituições de Pesquisa. São credenciadas ao Ministério da Educação e ao Ministério da Ciência e Tecnologia do Brasil e compõem o universo deste estudo. O Conselho Nacional das Fundações Estaduais de Amparo à Pesquisa (CONFAP) é uma organização sem fins lucrativos que tem por objetivo melhor articular os interesses das agências estaduais de fomento à pesquisa. Criado oficialmente em 2006, o Conselho agrega FAP de 24 Estados, além do Distrito Federal (BRASÍLIA, 2014).

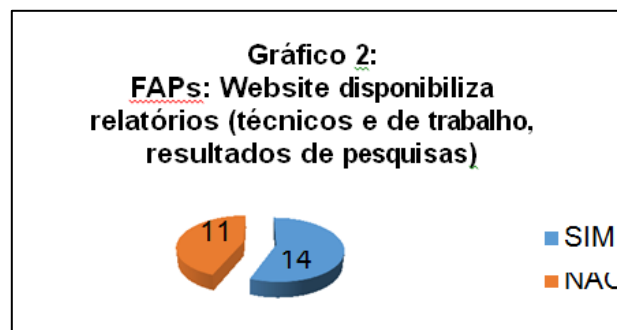
A presente pesquisa se configura com abordagem descritiva, do tipo aplicada e de natureza quantitativa. Utilizou-se da técnica do levantamento de dados, realizado ao nível nacional, nos vinte e cinco (25) websites das Fundações Brasileiras de Apoio à Pesquisa, que configuram a amostra do estudo.

Cenário nacional de acesso a dados de pesquisas financiadas pelas Fundações de Amparo a Pesquisas

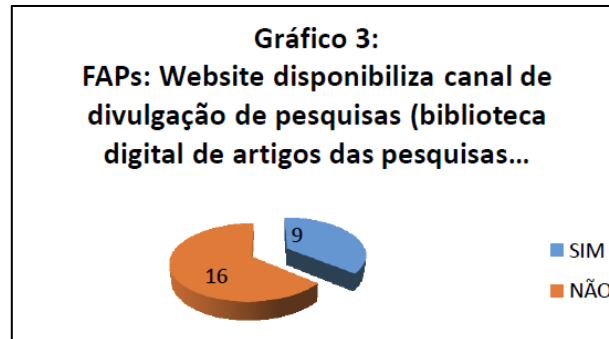
Os resultados de realce estão apresentados a seguir. Inicialmente, verificou-se se os websites das FAP disponibilizam listagens de projetos apoiados. O gráfico 1 ilustra que uma quantidade expressiva de FAP ainda não disponibiliza essas listagens. Tal resultado aponta que 52% das FAP não divulga os projetos realizados, com isso, observa-se a falta de visibilidade dos projetos enquanto estão sendo executados. Alerta-se para a importância de essas informações serem postadas nos websites das Fundações, que todavia não o fazem, para que o acesso aos dados possa mostrar o panorama momentâneo e, com isso, facilitar o desenvolvimento de outros projetos e redirecionar ações similares.



Outro item verificado foi a existência de relatórios técnicos ou de ações, bem como de resultados de pesquisas financiadas. O gráfico 2 ilustra que 56% das FAP disponibiliza este tipo de informação. Todavia, indica-se que as demais FAP (44% da amostra) passem a divulgar os relatórios e resultados, na íntegra com o propósito de apoiar estudos em andamento além de fomentar novas pesquisas.

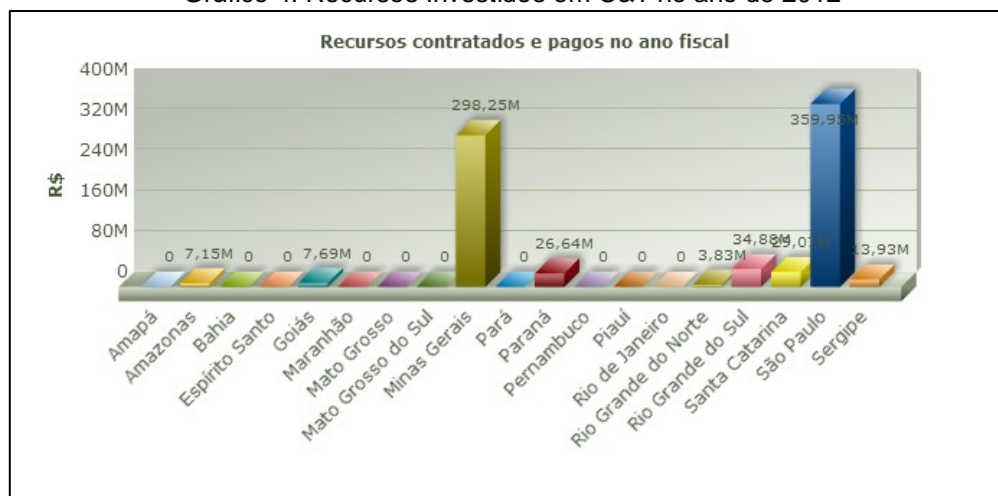


O resultado do levantamento deste item, aponta que 64% das FAP não possuem revistas eletrônicas para a divulgação das pesquisas fomentadas, além de bibliotecas digitais com os artigos que trazem os resultados das investigações financiadas, como pode ser observado no gráfico 3. Trata-se de um quantitativo interessante e pode-se inferir que ocorra incentivo para a visibilidade dos resultados dos estudos empreendidos. Advoga-se, também, pela participação das demais FAP nesse sentido.



A seguir o gráfico 4 aponta os valores de recursos investidos em Ciência e Tecnologia em todas as FAP brasileiras, no ano de 2012. Observa-se maior aporte nos Estados de São Paulo e Minas Gerais; seguidos dos Estados do Rio Grande do Sul e Paraná, isto é, incentivos de relevo nas regiões: sudeste e sul. Sugere-se estudo aprofundado para obtenção de análises detalhadas acerca dos fatores intervenientes em tais resultados.

Gráfico 4: Recursos Investidos em C&T no ano de 2012





O Livro Azul, resultado da 4ª Conferência Nacional de Ciência, Tecnologia e Inovação para o Desenvolvimento Sustentável, sob a coordenação do Ministério da Ciência e Tecnologia, na cidade de Brasília traz como recomendação para o avanço nacional da ciência, no tocante à distribuição de recursos:

É necessário garantir o atendimento às demandas mais sensíveis às assimetrias intra e inter-regionais, concomitantemente ao processo de indução no financiamento da pesquisa científica e tecnológica. Cabe desenvolver mecanismos de descentralização dos recursos para o fortalecimento dos sistemas regionais de C,T&I. Deve ser buscado um melhor entendimento das bases sobre as quais uma agenda de desconcentração possa ser ampliada (2010,p.68).

CONSIDERAÇÕES FINAIS

Sabe-se que, ao longo do tempo, a ciência teve como foco a busca por tecnologias que pudessem decifrar e dimensionar o mundo. Como consequência desta busca incessante, tem-se grande volume de dados gerados em laboratórios e institutos de pesquisa abastecendo planilhas infundáveis. Diante disto, pode-se afirmar que a ciência vive hoje um paradigma que vem alterando os rumos da pesquisa científica, a E-Science. De fato, a consulta aos dados de pesquisa ora concluídas, por outras investigações em elaboração, ou ainda, em vias de serem realizadas depende de uma efetiva gestão desses dados para que possam ser armazenados em bases confiáveis garantindo a recuperação ao longo do tempo. Na tentativa de propor ações para gerir dados de pesquisa, tem-se a curadoria digital que se relaciona com o modo de disponibilizá-los em formato digital para acesso futuro. A ideia é tornar esses dados acessíveis, interpretáveis, reutilizáveis e confiáveis para pesquisas ainda não realizadas.

Sabe-se que ao solicitar fomento para pesquisas científicas, exige-se a apresentação de projetos iniciais e, ao final, na maioria das vezes os resultados obtidos. Uma vez que as agências de fomento já estão de posse destes dados, torna-se mais fácil a implementação de repositórios específicos de dados científicos de pesquisas financiadas.

Ao concluir o presente estudo, observou-se que os resultados obtidos sugerem certa visibilidade dos resultados de pesquisas fomentadas no país com atenção especial voltada para a apresentação de como as agências de fomento têm empregado os recursos disponíveis. Contudo deve-se empreender melhorias quanto ao tratamento de dados digitais, bem como a disponibilização de informações e conteúdos científicos. Sugere-se o planejamento, a implantação e(ou) a implementação de sistemas de gestão desses ativos, por meio de repositórios digitais, institucionais e(ou) temáticos, no âmbito nacional. Essa iniciativa permitiria facilitar o acesso, o uso e a reutilização de informações e de conteúdos de realce para o avanço de pesquisas futuras. Ademais, a ideia de curadoria digital se amplia além do controle do repositório que arquiva os recursos e envolve a atenção do criador do conteúdo e dos usuários futuros, indicando-se, portanto, a garantia da preservação e da custódia em ambiente digital, para uso e usufruto contínuo de pesquisadores e interessados.



REFERÊNCIAS

ABOTT, Mark R. Um novo camino para a ciencia? In: HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin. *The Fourth Paradigm: Data-Intensive Scientific Discovery*, 2009. Disponível em: <http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_jim_gray_transcript.pdf>. Acesso em: 20 jul. 2014.

CONSELHO NACIONAL DAS FUNDAÇÕES ESTADUAIS DE AMPARO À PESQUISA (CONFAP). Disponível em: <<http://confap.org.br/news/>>. Acesso em: 08 jul. 2014.

CORRÊA, Elisabeth Saad; BEROCCHI, Daniela. O papel do comunicador num cenário de curadoria algorítmica de informação. In: CORRÊA, Elisabeth Nicolau Saad (Org.). *Curadoria digital e o campo da comunicação*. São Paulo: ECA/USP, 2012. p.22-29. Disponível em: <http://issuu.com/grupo-ecausp.com/docs/ebook_curadoria_digital_usp/57#download>. Acesso em: 20 mai. 2014.

COSTA, Sely Maria de Souza. O novo papel das tecnologias digitais na comunicação científica. In: MARCONDES, Carlos H. et al. (Org.). *Bibliotecas digitais: saberes e práticas*. Brasília: IBICT, 2005. p. 167-185.



Acesso Aberto

Legislação, impacto e visibilidade, modelos sustentáveis para publicações, políticas, acesso aberto e educação superior, métricas de impacto, novas tecnologias e ferramentas



ACESSO ABERTO NO CONTEXTO DOS REPOSITÓRIOS DIGITAIS EDUCACIONAIS: O CASO DA REDE UNA-SUS

*Fernanda de Souza Monteiro¹, Aline dos Santos Jacob², Soraya Alves Lacerda³,
Vinícius de Araújo Oliveira⁴*

¹Professora Adjunta da Universidade de Brasília, Doutora em Ciência da Informação pela Universidade de Brasília E-mail: fernanda.s.monteiro@gmail.com

²Bolsista da Universidade Aberta do SUS, Bacharel em Biblioteconomia pela Universidade de Brasília. Pós-graduanda em Informação Científica e Tecnológica em Saúde pela Fiocruz e-mail: alinejacob@unasus.gov.br

³Bolsista da Universidade Aberta do SUS, Bacharel em Biblioteconomia, Pós-Graduação em Análise de Sistemas pelo UnB/GFI Informática. Especialista em Marketing de Serviços e Inteligência Competitiva, pela FGV E-mail: sorayalacerda@unasus.gov.br

⁴Médico formado pela Universidade Federal de Minas Gerais, Mestre em Saúde Pública pela Universidade Federal de Minas Gerais, Coordenador Técnico na Secretaria Executiva da Universidade Aberta do SUS. E-mail: vinciusoliveira@unasus.gov.br

RESUMO

Estratégias de ensino-aprendizagem na área de saúde, fundamentais para a melhora da eficácia de todo sistema de saúde, requerem investimentos e esforço na formação e qualificação dos trabalhadores. Por esse motivo, foi criado em dezembro de 2010 o Sistema Universidade Aberta do SUS (UNA-SUS), por meio do Decreto nº 7.385/2010. A UNA-SUS é composta por três elementos estruturantes: uma rede de universidade que oferece cursos para qualificação dos trabalhadores de saúde (Rede UNA-SUS), um sistema de informações sobre os trabalhadores de saúde do Brasil (Plataforma Arouca), e o Acervo de Recursos Educacionais em Saúde (ARES). Com o intuito de estabelecer padrões mínimos de negociação de direitos autorais que viabilizem o acesso aberto a objetos de aprendizagem publicados no repositório educacional digital da UNA-SUS, o ARES, foi iniciada a implantação de uma Política de Acesso Aberto da UNA-SUS. Essa Política estabelece diretrizes para garantir o acesso, utilização e reutilização de conteúdos educacionais, respeitando os direitos morais do autor. Para tanto, foi realizada pesquisa bibliográfica sobre os temas relacionados ao acesso aberto e direitos autorais, seguida do levantamento de necessidades da Rede UNA-SUS. A definição de instrumentos legais para assegurar o processo de gestão dos direitos autorais foi feita com o apoio de um consultor jurídico, validação por um grupo de trabalho constituído no âmbito da Rede e posterior implementação dos elementos que constituem a Política de Acesso Aberto da UNA-SUS, tais quais: termos de cessão de direitos autorais, autorizações para o uso de imagem e voz, termos de uso e licença de distribuição dos recursos educacionais publicados no repositório digital educacional. A implantação de uma Política de Acesso aberto reforça o objetivo da UNA-SUS, qual



seja o de fomentar e apoiar a disseminação de meios e tecnologias de informação e comunicação que possibilitem ampliar a escala e o alcance das atividades educativas para os profissionais da saúde.

Palavras-chave: Política de acesso aberto. Objetos de aprendizagem. Direitos autorais.

ABSTRACT

With the purpose of establishing minimum standards of copyright negotiation that enables open access to learning objects published in the repository of educational digital UNA-SUS - Acervo de Recursos Educacionais em Saúde – ARES -, the implementation of the Open Access Policy for UNA-SUS was initiated. To do so, a bibliographical research related to open access and copyright was taken, followed by the needs assessment of UNA-SUS Network. The definition of legal instruments to ensure the process of copyright management was made, with the assistance of a legal advisor, and validated by a working group set up under the Network scope, and by the subsequent implementation of the elements that constitute the core of UNA-SUS Open Access Policy.

Keywords: Open Access policy. Learning objects. Copyrights.

INTRODUÇÃO

O Brasil vive um momento de grande demanda por serviços de saúde de qualidade. Como o setor de saúde é muito dependente de recursos humanos, a demanda por qualificação e educação permanente do pessoal de saúde é uma questão a ser enfrentada. Esse artigo discute como o acesso aberto tem sido uma peça fundamental no enfrentamento dessa questão, principalmente no contexto de disseminação de conteúdos em repositórios digitais educacionais.

A ampliação de cobertura de serviços preventivos baseados em agentes de saúde, enfermeiras clínicas e médicos de família melhora a eficácia de todo sistema de saúde, mas inevitavelmente requer investimentos e demanda esforço de formação e qualificação dos trabalhadores. Por esse motivo, o Brasil criou em dezembro de 2010 o Sistema Universidade Aberta do SUS (UNA-SUS), por meio do Decreto nº 7.385/2010. A UNA-SUS é composta por três elementos estruturantes: uma rede de universidades que oferecem cursos livres para qualificação dos trabalhadores de saúde - Rede UNA-SUS -, um sistema de informações sobre os trabalhadores de saúde do Brasil - Plataforma Arouca -, e o Acervo de Recursos Educacionais em Saúde – ARES. Os conteúdos educacionais produzidos pela UNA-SUS são disponibilizados no ARES e, assim, o investimento realizado pelo Ministério da Saúde na produção de cursos para educação a distância é transformado em patrimônio público, em acesso aberto na Internet (BRASIL, 2011). Isso impede que os conteúdos educacionais, produto de



financiamento público para oferta de cursos, fiquem restrito às instituições ou grupos. Assim, espera-se que o ARES contribua para uma progressiva redução nos custos por aluno em ofertas educacionais, uma vez que menor esforço precisa ser despendido em novas produções.

O principal beneficiário do ARES é o estudante ou trabalhador em saúde, que passa a ter acesso a esse conteúdo para seus estudos independentes, ou de forma complementar a outros processos educacionais. Para atingir esse propósito o ARES é federativo, fomenta o acesso aberto, garante a qualidade do que é disponibilizado e é baseado em padrões. Um acervo federativo é mantido por diversas instituições, cooperando com base em diretrizes comuns. O acesso aberto possibilita ao interessado localizar e acessar o conteúdo na Internet, assim como utilizá-lo para os fins adequados. A garantia de qualidade é um diferencial importante do ARES quando comparado com outras coleções disponíveis na Internet. Nenhum recurso é publicado sem que uma instituição educacional credenciada ateste que ele passou por processos criteriosos de validação em três dimensões: científica, pertinência ao contexto do público-alvo e adequação midiático-pedagógica. Por fim, a UNA-SUS estabelece recomendações de desenvolvimento e padrões para os conteúdos disponibilizados para que sejam compatíveis com equipamentos e contextos distintos de utilização.

No que diz respeito, especificamente, ao acesso aberto, a UNA-SUS identificou a necessidade de gerir adequadamente os direitos autorais ainda na produção de conteúdos, a fim de garantir sua disponibilização no ARES para posterior utilização. Assim, foi iniciado um estudo para definir uma política de acesso aberto com elementos suficientes à orientação da Rede. As seções seguintes abordam essa temática e elencam etapas e resultados deste estudo, em fase de conclusão.

ACESSO ABERTO E DIREITOS AUTORAIS

A Budapest Open Access Initiative, em 2001, definiu como acesso aberto a literatura disponível livre e publicamente na Internet, para qualquer usuário buscar, ler, copiar, imprimir, distribuir, etc. Podendo o conteúdo ser utilizado com qualquer propósito legal, sem barreiras financeiras ou técnicas, além daquelas indissociáveis do próprio acesso à Internet (BOAI, 2001). Os direitos autorais são o conjunto de prerrogativas que a lei reconhece a todo criador intelectual sobre suas produções literárias, artísticas ou científicas – com alguma originalidade, de ordem moral e patrimonial, durante toda a sua vida, com o acréscimo, para os sucessores indicados na lei, do prazo por ela fixado (CARBONI, 2003). Neste contexto, a função dos direitos autorais é preservar os direitos do autor de utilização de sua criação, resguardando sua integridade e estabelecendo os limites da distribuição e uso por terceiros.

Uma vez que as obras disponibilizadas em acesso aberto devem ter a identificação de autoria e descrição de conteúdo, as consequências são a visibilidade do autor e sua criação, além da promoção da produção científica e a citação.



É notável que o ideal de acesso aberto não tem o intuito de ferir, de nenhuma maneira, os direitos autorais. Há o entendimento de que a produção de conhecimento esteja acessível, no cumprimento de sua função social, sempre respeitando estes direitos. Além disso, a visibilidade do autor aumenta na medida em que se dá acesso ao conhecimento.

O acesso democrático ao conhecimento é um dos objetivos das instituições públicas de ensino, como aquelas que compõem a Rede UNA-SUS. Nesta perspectiva, a gestão dos direitos autorais irá garantir à Rede autonomia para estabelecer os termos de uso do que produzem, permitindo a reprodução, exibição, execução, apresentação, arquivamento, etc. por qualquer pessoa, física ou jurídica, pública ou privada, exclusivamente para fins privados, pedagógicos, didáticos, educacionais, de pesquisa, científicos e informativos, resguardando os direitos morais do autor.

Para tanto, a gestão dos direitos autorais inicia-se na produção de recursos educacionais, com a negociação prévia com autores e obtenção das devidas cessões. A falta de conhecimento e padronização das estratégias a serem adotadas neste momento acarretam problemas na atribuição de autoria, além dos que se referem a autorização para disseminação em repositórios.

PRODUÇÃO DE OBJETOS DE APRENDIZAGEM

A educação à distância, mediada pelas tecnologias de informação e comunicação (TICs), propicia o desenvolvimento de novas práticas pedagógicas e a utilização de uma diversidade de documentos digitais. Neste contexto, a produção de Objetos de Aprendizagem (OAs) é elemento fundamental para a composição de ações de ensino-aprendizagem. Esses OAs são conteúdos como textos, vídeos, animações, simulações, etc., estruturados para promover a estratégia educacional, ou seja, são unidades educacionais de pequena dimensão, desenhadas e desenvolvidas para fomentar sua utilização, eventualmente, em mais de um curso ou em contextos diferenciados, passíveis de combinação e/ou articulação uns com os outros de modo a formar unidades mais complexas e extensas (PIMENTA, BATISTA, 2004, p. 102).

Na produção de OAs são envolvidos diferentes perfis profissionais, tais como, desenhistas instrucionais, desenvolvedores, especialistas de domínio, fotógrafo, etc. A ação sistemática e conjunta desses diversos atores gera uma cadeia complexa de direitos autorais e titularidade sobre os OAs.

A equipe multidisciplinar envolvida na produção trabalha de maneira coordenada para a concepção final do OA, porém, possui atribuições e produtos individuais, concluídos em momentos distintos. O autor é definido por sua contribuição intelectual e/ou artística na composição de uma obra, os demais envolvidos são compreendidos como colaboradores, não cabendo a estes direitos autorais. A obra pode ser em coautoria, quando é criada por dois ou mais autores, ou coletiva, criada por iniciativa, organização e responsabilidade de uma pessoa física ou jurídica, que a publica sob seu nome ou marca e que é constituída pela participação de diferentes autores, cujas contribuições se fundem numa criação autônoma (BRASIL, 1998).



Nota-se que os problemas relacionados aos direitos autorais de objetos de aprendizagem decorrem, portanto, da complexidade no processo de produção. Por isso, a definição de requisitos de produção que priorizem a gestão e a definição de instrumentos legais de cessão de direitos patrimoniais é fundamental para garantir a disseminação dos recursos educacionais em acesso aberto.

As instituições da Rede UNA-SUS coordenam sua produção e são titulares dos direitos autorais dos OAs, que são obras coletivas. Cada autor desenvolve conteúdos específicos que compõem o OA, e cede os direitos patrimoniais para a instituição, mas mantém seus direitos morais. Assim, os autores têm o direito de serem reconhecidos pela sua produção, ainda que ela seja parte integrante de uma obra maior, ou OA final. Vale ressaltar, que a Rede não exige uma cessão exclusiva, mantendo o direito dos autores utilizarem seus conteúdos de maneira irrestrita, inclusive para venda.

Os recursos educacionais produzidos pela Rede UNA-SUS são disponibilizados em seu repositório digital educacional, o Acervo de Recursos Educacionais em Saúde – ARES. A disseminação neste ambiente contribui para a diminuição de custos de produção, ao favorecer a reutilização de recursos.

DISSEMINAÇÃO DE OBJETOS DE APRENDIZAGEM

De maneira geral, a *Word Wide Web (Web)* tem favorecido a publicação e disseminação de conteúdos informacionais. Esses conteúdos podem ser acessados em *blogs, sites*, portais, entre outras fontes de informação. Nesse contexto, crescem as soluções para o agrupamento de documentos que apresentam conteúdos com objetivos e/ou temas comuns. Uma dessas soluções é o Repositório Digital (RD), que permite, além do armazenamento, a descrição e busca de informação em formatos digitais.

O RD surge no contexto acadêmico, a partir da iniciativa dos arquivos abertos, com a pretensão de disponibilizar uma variedade de documentos não publicados tradicionalmente por editoras e revistas científicas (HARNAD, 2001). Esses RDs possuem uma concepção mais ampla que a das bibliotecas digitais no sentido de permitir a disseminação de documentos em versões preliminares para apreciação de pares e podem ter diferentes definições que variam de acordo com seu contexto de desenvolvimento e objetivo. A definição do *Digital Repositories JISC3* (HAYES, 2005) é ampla e adequada aos propósitos desse projeto ao afirmar que um repositório digital é aquele onde os conteúdos informacionais digitais podem ser pesquisados e recuperados para uso posterior.

Os RDs apresentam estrutura e características próprias, capazes de permitir a um usuário comum a recuperação e o acesso a documentos de seu interesse. Geralmente estão sob a responsabilidade técnica e administrativa de uma instituição ou organismo (CAFÉ, 2003). Em um RD destacam-se as seguintes características gerais (HEERY; ANDERSON, 2005 apud MARTINS; RODRIGUES; NUNES, 2008):



- a) Suporta a submissão de documentos pelo autor, profissional de informação ou responsável pelo repositório;
- b) Disponibiliza documentos acrescidos de seus metadados descritivos;
- c) Oferta serviços de busca, localização e acesso aos documentos;
- d) Garante a atualização e disponibilidade do conteúdo a partir de uma gestão sustentável.

Um RD com conteúdo e funcionalidades focadas na educação é chamado Repositório Digital Educacional (RDE), diferenciando-o dos demais. A preocupação de seus mantenedores não é disponibilizar documentos que sirvam apenas de referência para pesquisas, levantamentos bibliográficos e subsídios teóricos, como a maioria dos RDs acadêmicos. Um RDE deve disponibilizar documentos que possam ser incorporados e utilizados diretamente no processo de ensino-aprendizagem.

Para isso, além das características gerais elencadas para o RD, o RDE deve apresentar uma arquitetura adequada, dispor de informações descritivas suficientes à seleção criteriosa de documentos, especialmente relacionadas aos aspectos educacionais. Também é recomendável que o RDE possa interagir com ambientes virtuais de aprendizagem possibilitando, por exemplo, a exportação e importação de seus documentos (MIRANDA, 2004; NASCIMENTO, 2010; MCLAREN, 2004).

Em RDEs pode-se disponibilizar uma grande variedade de recursos de ensino-aprendizagem, dentre eles aqueles definidos como Objetos de Aprendizagem (OAs).

Os OAs são disponibilizados em RDEs e podem ser combinados e recombinados compondo novos e mais complexos OAs. A organização e tratamento informacionais adequados para esses objetos são primordiais para garantir sua localização e utilização no momento desejado. Dada a diversidade dos OAs e dos seus usuários, os RDEs assumem fundamental importância na sua recuperação. Não é apenas um documento de referência o que se busca, é um objeto que condensa, aplica, demonstra, instrui, apresenta um conteúdo. Assim, o RDE possui um papel que transcende o de fonte de informação sobre determinado assunto e passa a fornecer conteúdos capazes de viabilizar a produção de uma estratégia educacional, permitindo a recuperação de OAs que podem, por exemplo, ser diretamente inseridos em um curso. O OA, dependendo de sua complexidade, assume parte substancial na composição de um curso, módulo de curso, unidade ou atividade de aprendizagem, podendo, até mesmo, ser um curso completo.

O RDE favorece a formação de coleções específicas, a organização, o compartilhamento, a recuperação e a reutilização de OAs. Isso reduz custos e amplia o público-alvo, propostas intrínsecas ao EAD. Assim, é necessário compreender não apenas a finalidade do RDE, mas a importância da sua alimentação, as diretrizes de desenvolvimento e a definição de padrões para submissão e acesso aos OAs.



PROCEDIMENTOS METODOLÓGICOS

Para atingir o objetivo de desenvolver uma política de acesso aberto capaz de subsidiar a produção de recursos educacionais da Rede UNA-SUS, bem como, garantir a ampla disseminação desses recursos em repositórios digitais e sua reutilização, foi realizada, inicialmente, uma pesquisa bibliográfica. A pesquisa bibliográfica permite fazer um histórico sobre o tema, atualizar-se e encontrar respostas para os problemas enfrentados no dia-a-dia das instituições que desejam tornar acessíveis materiais para serem utilizados por profissionais de saúde em estratégias de ensino-aprendizagem (AMARAL, 2007).

Nesse sentido, foram pesquisados documentos nacionais e internacionais que corroboraram com o entendimento da relação entre acesso aberto e direitos autorais, discutiram as especificidades da produção de objetos de aprendizagem e sua disseminação. O resultado desta etapa constituiu o referencial teórico, apresentado anteriormente. Esse referencial orientou a definição dos demais procedimentos metodológicos necessários para a elucidação de uma proposta para Gestão de Direitos autorais na Rede UNA-SUS. Os procedimentos metodológicos realizados são:

- a) Pesquisa bibliográfica sobre os temas relacionados;
- b) Levantamento de necessidades e equiparar procedimentos adotados pela Rede UNA-SUS para lidar com questões relacionadas a direitos autorais;
- c) Definição de instrumentos legais necessários à gestão de direitos autorais para que sejam propostos modelos com curadoria de especialista;
- d) Validação dos modelos propostos com a Rede UNA-SUS.

DESENVOLVIMENTO DA POLÍTICA DE ACESSO ABERTO DA UNA-SUS

Nesta seção são discutidos os resultados obtidos com a realização dos procedimentos metodológicos. Essa é uma pesquisa em fase de conclusão e, portanto, a validação de modelos propostos está em andamento.

A pesquisa bibliográfica, entre outras contribuições já mencionadas, identificou estratégias e diretrizes atualmente adotadas para a gestão de direitos autorais no contexto dos RDEs. A partir da compreensão de conceitos fundamentais elencados no referencial teórico, foi feito o levantamento das necessidades enfrentadas pela Rede UNA-SUS para garantir a disseminação e o acesso a sua produção de recursos educacionais, tanto por membros da própria Rede quanto por demais interessados.

O levantamento das necessidades foi realizado em oficinas presenciais para o alinhamento de estratégias para a Rede UNA-SUS como um todo. Essas oficinas presenciais ocorreram ao longo do ano de 2012, totalizando 3 (três) encontros principais em que os coordenadores das instituições que compõem a Rede tiveram oportunidade de manifestar livremente suas necessidades. Além disso, também foram realizadas entrevistas com pessoas envolvidas na produção de recursos educacionais,



a fim de facilitar a compreensão a respeito dos atores envolvidos no processo de produção, dos produtos parciais e finais entregues e dos documentos originados desde a contratação dos atores até a publicação dos recursos educacionais no ARES.

Esse levantamento possibilitou identificar e equiparar procedimentos adotados pela Rede UNA-SUS para lidar com questões relacionadas aos direitos autorais. Evidenciou-se que, com ou sem orientação de especialistas, as instituições já vinham resguardando alguns direitos de uso daquilo que era demandado e produzido. Ainda assim, no que concerne a gestão dos direitos autorais, observou-se que:

- a) É fundamental a orientação e uniformidade das instruções passadas às instituições para garantir a disseminação em repositórios, uso e reuso dos recursos educacionais;
- b) Há a necessidade de definição de um termo de cessão padrão para a Rede UNA-SUS, responsável pela produção, para o Ministério da Saúde, como principal demandante, e conseqüentemente para os usuários em geral, que irão contribuir com o acesso aberto a essas informações;
- c) É essencial a delimitação das condições de uso e reuso dos objetos disponibilizados no ARES, por meio da definição de termos de uso, a serem observadas pelo usuário final;
- d) Há dificuldade em acompanhar a gestão dos direitos autorais das obras, o que gera demora na publicação de recursos no ARES e problemas entre os autores de recursos e as universidades.

Para a criação de instrumentos que pudessem atender a estas necessidades de forma a garantir a política de acesso aberto ao ARES, A UNA-SUS contou com a colaboração de um Consultor, *expert* em direitos autorais, o qual foi convidado a traçar um panorama dos direitos autorais no âmbito do Sistema UNA-SUS. O trabalho deste consultor foi apoiado, consolidado e validado pelo Grupo de Trabalho formado por coordenadores da Rede UNA-SUS e pela Equipe de Ciência da Informação da SE/UNA-SUS, responsável pela gestão do ARES.

Após a realização de vários encontros, no âmbito da SE/UNA-SUS, entre coordenadores, o consultor em direitos autorais e as equipes relacionadas, culminou-se na produção de 4 (quatro) instrumentos mínimos para a gestão dos direitos autorais. São eles:

- a) Termo de cessão padrão (caráter não exclusivo e não comercial);
- b) Termos de uso para o usuário final dos recursos;
- c) Sistema de geração padronizada e de gestão dos termos de cessões;
- d) Lista de perguntas frequentes elucidação de dúvidas.



Termo de Cessão Padrão UNA-SUS

O termo de cessão (TC) padrão tem o papel de desenhar a via mestra de trânsito dos direitos autorais nas relações entre o autor, as universidades e a UNA-SUS. Ele foi construído visando definir um documento a ser adotado como modelo por todas as instituições da Rede UNA-SUS de modo a facilitar as questões relativas a negociação de direitos autorais nas universidades, abrangendo não somente os direitos da UNA-SUS, como também os do usuário final que fará uso dos recursos educacionais publicados no ARES (Anexo A).

O TC padrão também será anexado aos termos de contratação e de convênios, a fim de garantir, logo no início desta relação de contrapartida, a adequação às diretrizes de direitos autorais da UNA-SUS.

Após estudos do fluxo de produção dos recursos educacionais pelas instituições participantes da Rede, do funcionamento de convênios e termos de cooperação, bem como das características gerais do ARES, chegou-se à conclusão de que o TC teria como padrão caráter não exclusivo e não comercial, pois assim assegura os usos de interesse da UNA-SUS no âmbito do ARES, resguardando todos os direitos morais do autor, e garantindo, ainda, uma maior adesão por parte destes e das Instituições participantes. Cessões com caráter comercial e/ou exclusivo e suas diferentes combinações, seriam tratadas como exceções, caso a caso e, por conseguinte, consideradas variantes deste TC padrão. Estas variantes seriam utilizadas, excepcionalmente, quando houver interesse em comercialização do material, a posteriori.

No caso da cessão padrão o autor cede e transfere ao Ministério da Saúde, à UNA-SUS, e à respectiva universidade produtora dos recursos, os direitos autorais do recurso educacional e das suas partes constitutivas. Com isso, os mesmos direitos de uso (patrimoniais), não comerciais e não exclusivos, são conferidos a todos, preservados os direitos morais do autor.

Termos de Uso

Com o TC padrão definido, coube então definir em que termos seria permitido o uso e reuso dos objetos disponibilizados no ARES. Os Termos de Uso (TU) de um Repositório definem as regras e os limites de uso para seus recursos. No caso do ARES a definição destes limites inexistiam, visto que os instrumentos anteriormente utilizados não regulavam todas as possibilidades de uso do(s) recurso(s) por seu usuário final, tão pouco as responsabilidades de cada ator nesta relação.

São usos autorizados para os recursos do ARES: reprodução, exibição, execução, declamação, exposição, apresentação, *download*, arquivamento por qualquer pessoa, física ou jurídica, pública ou privada, exclusivamente para fins privados, pedagógicos, didáticos, educacionais, de pesquisa, científicos, informativos.



Todos os usuários terão os mesmos direitos dispostos na licença da UNA-SUS, exceto os de criação de obras derivadas, tais como: disponibilização pública; tradução para qualquer idioma; inclusão em banco de dados; inclusão em novas obras ou coletâneas; inclusão formal em material didático ou de curso. Para estes, uma autorização prévia e expressa deverá ser solicitada à UNA-SUS.

Sistema Gestão dos Termos de cessões

Atualmente, a negociação dos direitos autorais referentes a cursos e recursos produzidos no âmbito do Sistema UNA-SUS, é de responsabilidade das instituições que compõem a rede, e não segue um padrão específico em todas elas. Por conseguinte, há uma dificuldade em acompanhar a gestão dos direitos autorais das obras, o que gera demora na publicação de recursos no ARES e problemas entre os autores de recursos e as Instituições de Ensino Superior.

Verificou-se, então, a necessidade de implantação de uma ferramenta que padronizasse a geração dos TCs - bem como suas eventuais variáveis -, permitindo a agregação dos seguintes benefícios:

- a) Compartilhamento facilitado de modelos de TCs padrão para uso por toda Rede;
- b) Fácil identificação e seleção das possíveis variáveis, a partir do TC padrão;
- c) Simplificação da geração dos TCs, permitindo agilizar a negociação dos direitos autorais, no tempo devido (antes da produção dos recursos);
- d) Garantir a gestão do fluxo cessão dos direitos autorais, permitindo o acompanhamento de entregas e o arquivamento das mesmas em um único ambiente.

O objetivo do sistema é utilizar o texto do TC padrão como ponto de partida - cujo escopo é de cessão não comercial, não exclusiva e sem ônus -, e a partir dele limitar os campos de preenchimento, controlando as opções que favorecem a geração de variações deste TC padrão. Assim, uma vez preenchidos os campos, e feitas as escolhas das variáveis, o sistema gerará um TC pronto a ser assinado.

Além das características já atribuídas ao TC padrão, são variações contempladas para geração no sistema:

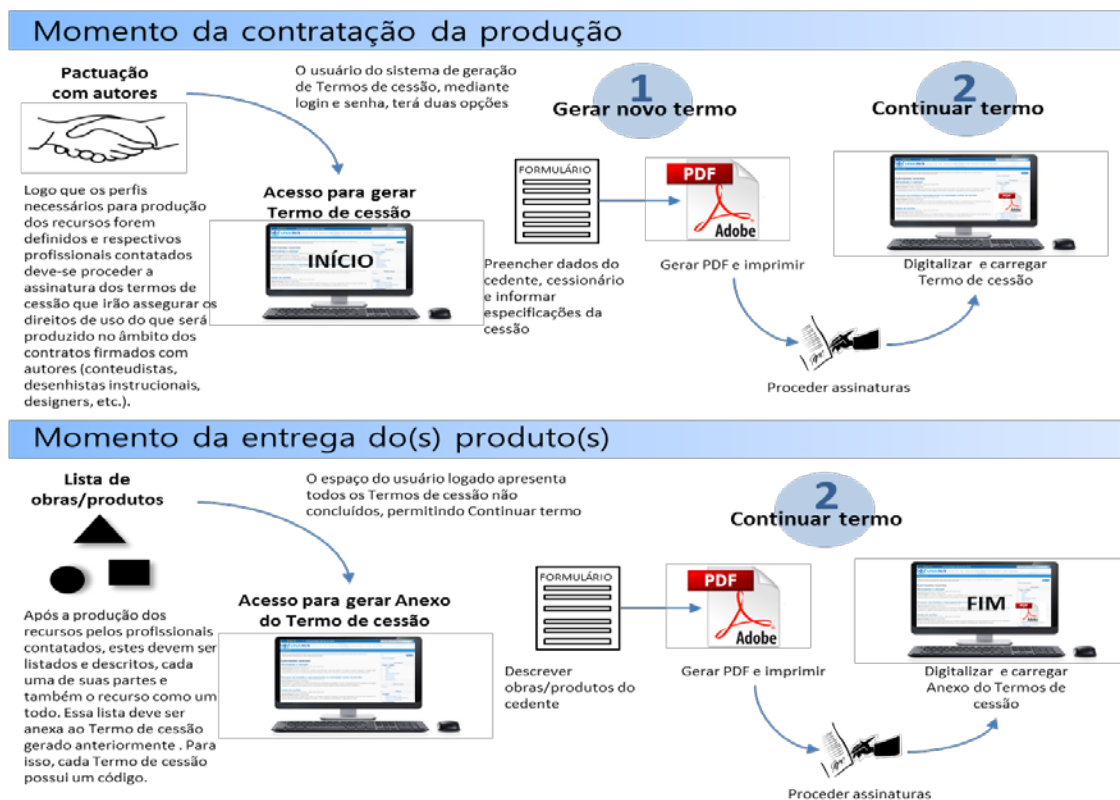
- a) Termo de cessão não exclusiva e comercial;
- b) Termo de cessão exclusiva e não comercial;
- c) Termo de cessão exclusiva e comercial.

Para cada TC gerado teremos um “número de controle” que será usado, também, para gerar automaticamente a identificação do cabeçalho do(s) anexo(s) onde serão listadas todas as obras produzidas sob as regras daquele TC. Este número identificará, ainda, o armazenamento e o controle da cópia digitalizada dos TCs e seus anexos, e será usado para vincular a cessão à descrição de seus recursos correspondentes no ARES.

O sistema contempla também a preservação dos TCs rubricados, permitindo o *upload* dos documentos (TCs e anexos) após recolhidas todas as assinaturas.

O fluxo geral do Sistema de Gestão de TCs (Figura 1) tem seu início na pactuação entre os atores da produção e a geração do TC, passando pela compilação dos produtos entregues nos anexos, e finalizando com o arquivamento dos documentos assinados.

Figura 1 - Fluxo do Sistema de Gestão de Termos de Cessão





Perguntas Frequentes

Além da elaboração dos documentos que irão subsidiar a implantação da política, foi identificada a necessidade de elaboração de um guia com respostas às perguntas mais frequentes sobre direito autoral, abordando seus principais conceitos.

A lista de Perguntas Frequentes foi criada a fim elucidar dúvidas, nivelar conhecimento e, com isso, auxiliar as universidades na adesão à nova política.

A intenção é que a lista de perguntas frequentes seja um instrumento orgânico, podendo ser atualizada de tempos em tempos, retro alimentadas por questões manifestadas comumente pelo canal de suporte e atendimento do ARES.

Estratégias de Implementação

Após as etapas colaborativas de criação e validação dos produtos iniciais, a UNA-SUS irá implementar estes instrumentos. Para isso, são propostos os seguintes passos, que constituem a estratégia de implementação:

- a) Desenvolvimento e implantação do sistema de gestão de termos de cessão;
- b) Criação de manual descrevendo o fluxo da gestão dos direitos autorais;
- c) Criação de um canal de comunicação para sanar dúvidas e retroalimentar a lista de pergunta frequentes.
- d) Desenvolvimento e oferta Curso sobre Direitos autorais, na modalidade a distância, a ser ofertado para a Rede UNA-SUS.

Gestão de Direitos Autorais na Rede UNA-SUS

Todo este processo de construção de uma Política de Acesso Aberto da UNA-SUS busca exatamente estabelecer parâmetros mínimos para a negociação e gestão dos direitos autorais referentes aos recursos educacionais produzidos no âmbito de sua Rede.

Esta iniciativa garantirá que as instituições da Rede, bem como o Ministério da Saúde e a própria UNA-SUS, possam utilizar, reutilizar e principalmente disponibilizar esses recursos em esferas educacionais diversas, desempenhando o papel real do Acesso Aberto, mas resguardando ainda os direitos do autor.

Com esse primeiro passo, a UNA-SUS propõe soluções simples e inovadoras para lidar, de forma descomplicada, mas segura, com os direitos autorais, que perpassam de forma transversal etapas que vão desde a contratação dos autores até o momento em que os recursos educacionais são disponibilizados no ARES para seu usuário final.



Desta forma, a Política de Acesso Aberto da UNA-SUS nasce contemplando o compartilhamento com as instituições da Rede UNA-SUS dos modelos de termo de cessão de direitos autorais, a divulgação dos termos de uso no ARES, a oferta de um sistema para geração e controle dos TCs, além de oferecer treinamento e orientações sobre como proceder em casos específicos do dia-a-dia da produção de recursos educacionais.

CONSIDERAÇÕES FINAIS

A UNA-SUS avança mais um passo na direção de cumprir com seu papel de democratizar a educação ao disseminar sua produção de OAs, regida por uma política de acesso aberto. A conclusão deste estudo e os produtos resultantes dele reforçam este papel social ao buscar padrões e entendimentos comuns no âmbito de sua Rede, ao conseguir combinar nos instrumentos produzidos, a preservação dos direitos autorais, ao mesmo tempo em que estende, do Ministério da Saúde até o usuário final do ARES, a possibilidade real de utilização, reutilização, disseminação e compartilhamento dos recursos produzidos. Outro aspecto relevante do trabalho é poder, enfim, atender às demandas das instituições participantes da Rede UNA-SUS, ansiosas por orientações para a condução de suas negociações de direitos autorais, dessa vez com padrões desenhados para atender suas necessidades. O Sistema de Geração de Termos de Cessão com certeza será um divisor de águas neste processo, contribuindo para a manutenção destes padrões, para a agilidade na pactuação das negociações, para a gestão do seu *workflow*, e também para a transparência deste pacto ao disponibilizar os recursos educacionais no ARES.

REFERÊNCIAS

AMARAL, João J. F. **Como fazer uma pesquisa bibliográfica**. - Ceará: Universidade Federal do Ceará, 2007. 21 p. Disponível em: <<http://200.17.137.109:8081/xiscanoe/courses-1/mentoring/tutoring/Como%20fazer%20pesquisa%20bibliografica.pdf>> acesso em: jul. 2011.

BRASIL. Lei nº 9.610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação de direitos autorais e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, 20 fev. 1998. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/L9610.htm>. Acesso em: out. 2012.

BRASIL. Ministério da Saúde. Secretaria de Gestão do Trabalho e da Educação na Saúde. Universidade Aberta do Sistema Único de Saúde. **Política do acervo de recursos educacionais em saúde**. Responsabilidade técnica: Fernanda de Souza Monteiro e Aline Santos Jacob. – Brasília : UNA-SUS, 2011.



Budapest Open Access Initiative. Dez anos depois da Budapest Open Access Initiative: estabelecendo o Acesso Aberto como padrão. Budapest: BOAI, 2001. Disponível em: <<http://www.budapestopenaccessinitiative.org/boai-10-translations/portuguese>>. Acesso em: jul 2013.

CAFÉ, Lígia. et al. Repositórios institucionais: nova estratégia para publicação científica na Rede. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 26, 2003, Belo Horizonte. **Anais...** Belo Horizonte: INTERCOM, 2003. Disponível em: <http://www.intercom.org.br/papers/nacionais/2003/www/pdf/2003_ENDOCOM_TRABALHO_cafe.pdf >. Acesso em: out. 2011.

CARBONI, Guilherme C. Direitos do autor na multimídia. São Paulo: Quarter Latin, 2003. HARNAD, Stevan. et al. **The access/impact problem and the green and gold roads to open access.** 2001. Disponível em: <<http://www.ecs.soton.ac.uk/~harnad/Temp/impact.html>>. Acesso em: fev. 2011

HAYES, Helen. **Digital Repositories:** helping universities and colleges. JISC Briefing Paper – Higher Education Sector. Helen Hayes (Ed.). England: August 2005. Disponível em: <<http://archive.excellencegateway.org.uk/page.aspx?o=117227>>. Acesso em: fev. 2010.

MARTINS, Ana Bela; RODRIGUES, Eloy; NUNES, Manuela Barreto. Repositórios de informação e ambientes de aprendizagem: criação de espaços virtuais para a promoção da literacia e da responsabilidade social. **Rede de bibliotecas escolares**, newsletter n. 3, 2008.

MIRANDA, Raquel Melo de. **Groa:** um gerenciador de repositórios de objetos de aprendizagem. 2004. 80 f. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal do Rio Grande do Sul, Porto alegre, 2004.

NASCIMENTO, Anna Christina de Azevedo. Repositórios no contexto da EaD: o Banco Internacional de Objetos Educacionais. **Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, v. 15, n. 29, 2010.

MCLAREN, I. New trends in web-based learning: objects, repositories and learner engagement. **European Journal of Engineering Education**, v. 29, n. 1, p. 65-71, 2004.

PIMENTA, P.; BATISTA, A. A. **Das plataformas de e-Learning aos objetos de aprendizagem.** In: E-Learning para E-formadores. Minho: TecMinho, 2004. p. 97; 109



ANEXO A

CESSÃO NÃO EXCLUSIVA DE OBRAS AUTORAIS PARA FINS NÃO COMERCIAIS

(nome e qualificação – domicílio, nacionalidade, estado civil, identidade e CPF), doravante denominado(s) CEDENTE(s), por meio deste instrumento, CEDE e TRANSFERE, gratuitamente, ao MINISTÉRIO DA SAÚDE/UNIVERSIDADE ABERTA DO SISTEMA ÚNICO DE SAÚDE e UNIVERSIDADE X, doravante denominadas CESSIONÁRIAS, em caráter permanente, irrevogável e NÃO EXCLUSIVO, os direitos patrimoniais NÃO COMERCIAIS de utilização da(s) obra(s) artística(s) e/ou intelectuais(s) em anexo, das quais declara ser (co)autor(a) e titular, durante o prazo de duração dos direitos autorais, em qualquer idioma e em todos os países.

A cessão total não exclusiva, permanente e irrevogável dos direitos autorais patrimoniais não comerciais de utilização de que trata este documento inclui, exemplificativamente, os direitos de disponibilização e comunicação pública da(s) obra(s), em qualquer meio ou veículo – principalmente, mas não unicamente, em Repositórios Digitais –, os direitos de reprodução, exibição, execução, declamação, exposição, arquivamento, inclusão em banco de dados, preservação, difusão, distribuição, divulgação, empréstimo, tradução, inclusão em novas obras ou coletâneas, modificação e transformação da(s) obra(s), reutilização, edição, produção de material didático e cursos ou qualquer forma de utilização não comercial.

A cessão aqui especificada concede ao MINISTÉRIO DA SAÚDE/UNIVERSIDADE ABERTA DO SISTEMA ÚNICO DE SAÚDE e/ou à UNIVERSIDADE X, independentemente da vontade da outra CESSIONÁRIA, o direito de autorizar qualquer pessoa – física ou jurídica, pública ou privada, nacional ou estrangeira a acessar e utilizar as obras para fins não comerciais, nos termos deste instrumento.

Para fins deste instrumento, usos não comerciais são aqueles em que as obras são disponibilizadas gratuitamente, sem cobrança ao usuário e sem intuito de lucro direto por parte daquele que as disponibiliza.

A não exclusividade dos direitos incluídos no âmbito deste instrumento contratual significa que tanto o CEDENTE como qualquer das CESSIONÁRIAS, individualmente considerados, poderá exercê-los de forma independente de autorização ou comunicação, prévia ou futura, às demais partes integrantes desta Cessão.

Ficam reservados exclusivamente ao CEDENTE todos os direitos morais sobre as obras de sua autoria e/ou titularidade, assim como os usos comerciais da(s) obra(s) incluídas no âmbito deste instrumento.

O CEDENTE declara ainda que a obra é de sua criação original, responsabilizando-se integralmente pelo seu conteúdo e outros elementos que fazem parte da(s) obra(s), obrigando-se em indenizar terceiros por danos, bem como indenizar e ressarcir as CESSIONÁRIAS de eventuais despesas que vier a suportar, em razão de qualquer ofensa mencionada nesta cláusula, principalmente no que diz respeito a plágios e violações de direitos autorais.



A cessão dos direitos autorais de utilização da(s) obra(s) artística(s) e/ou intelectuais(s) listada(s) em anexo, conforme estabelecidos neste instrumento, será gratuita, não sendo devida qualquer remuneração, a qualquer título, ao(s) autor(es) e titular(es).

LOCAL, DATA, ASSINATURA



AS REVISTAS ACESSO ABERTO E AS POLÍTICAS DE DIREITOS AUTORAIS: O CASO DE REVISTAS NA PLATAFORMA SCIELO-BRASIL

Ariadne Chloe Mary Furnival¹; Bianca Miranda de Almeida²

¹ Professora Dra. Associada, Depto. Ciência da Informação, Universidade Federal de São Carlos, S.P., Brasil. ² Bibliotecária e Coordenadora do Sistema de Bibliotecas do Centro Universitário UniSEB, Riberão Preto, S.P, Brasil.

RESUMO

O acesso aberto se refere à disponibilidade *online*, gratuita e livre de restrições, da literatura científica. Um modo muito divulgado de disponibilizar tal literatura, sobretudo no Brasil, é através das revistas acesso aberto, a chamada “via dourada”, sendo que comparativamente, os repositórios digitais abertos, nos quais versões de artigos são depositados – a “via verde” – são ainda menos utilizados no Brasil. Especificamente, é reconhecido que a publicação em revistas acesso aberto tem crescido continuamente desde o início da plataforma SciELO no Brasil em 1997, e a ampla promoção do sistema OJS pelo IBICT. Apesar da larga divulgação mundial dos conceitos e definições de acesso aberto, e sobretudo em relação aos direitos autorais e autorização de reuso e arquivamento em repositórios, a hipótese aqui levantada é que ainda existem posições conflitantes nas políticas de algumas revistas acesso aberto brasileiras relativo a esses direitos autorais e os direitos dos leitores-usuários. Diante do colocado, o presente estudo teve como objetivo principal analisar e averiguar o conceito de acesso aberto explícita e implicitamente usado nas revistas levantadas, sinalizando, de certa forma, o grau de “abertura” destas, tal como definido recentemente no folheto “*HowOpenIsIt?*” publicado pela *Scholarly Publishing and Academic Resources Coalition* (SPARC), a *Public Library of Science* (PLoS) e a *Open Access Scholarly Publishing Association* (OASPA). Assim, foi realizada uma pesquisa documental, baseada em leitura de documentos, seguido por um levantamento não-exaustivo das políticas de direitos autorais de uma amostra composta por 117 periódicos correntes da área das Ciências Biológicas e Ciências de Saúde na plataforma SciELO-Brasil. Estes dados foram confrontados com dados cadastrais encontrados no *Directory of Open Access Journals* (DOAJ), e subsequente análise usando o *software* Microsoft Excel. Os resultados mostram que dos 117 periódicos analisados no SciELO, 106 estão indexados no DOAJ e portanto, são classificados como revistas acesso aberto de acordo com os critérios de inclusão no DOAJ. Embora todas estas usem licenças *Creative Commons* BY-NC, apenas 25 revistas explicitamente divulgam em seu *website* que são de acesso aberto. Por fim, houve um conjunto substancial de 74 revistas que, apesar de estarem indexadas no DOAJ, e carregarem a licença CC BY-NC, adotam como política autoral a cessão de direitos pelo autor, prática essa que, a nosso ver, mina a denominação da revista como sendo de fato acesso aberto, já que o autor não



será o detentor dos direitos do artigo de sua autoria, implicando que possivelmente sofreria restrições na sua liberdade de auto-arquivamento do artigo num repositório institucional, por exemplo. Os resultados demonstram que há, de fato, uma variedade de apropriações e significações do termo “acesso aberto” pelas revistas analisadas, que se manifestam na redação da política dos direitos autorais da revista. Tal variedade dá espaço para confusões entre autores e leitores no que se refere aos seus direitos de reuso e disseminação, direitos estes intrínsecos às definições “clássicas” de acesso aberto.

Palavras-chave: Revistas acesso aberto; Direitos autorais; *copyright*; Direitos de reuso; SciELO; DOAJ.

INTRODUÇÃO

O acesso aberto se refere à modalidade de publicação que visa disponibilizar a literatura científica *online*, gratuitamente e livre de restrições. Um dos motivos por trás da ampla promoção dessa forma de publicação científica é que a pesquisa básica e aplicada é majoritariamente financiada pelos governos com a expectativa de que, ao compartilhar e difundir as novas ideias e descobertas resultantes, a ciência e tecnologia progredirá de modo mais rápido e eficiente, sem duplicações desnecessárias, que, por sua vez, melhorará a vida dos indivíduos e o bem estar da sociedade como um todo (GUIBAULT, 2011; RODRIGUES, 2004). Alguns autores observam a ironia de que justamente no momento em que as tecnologias de informação e comunicação (TIC) permitem maior acesso e uso da produção científica e criativa, existe simultaneamente um movimento de fortalecer o arcabouço de restrições que trancam e protegem o conteúdo digital em formas que nunca teríamos imaginado com material impresso (MONBIOT, 2011). Situação essa que levou a Diretora da *British Library* a perguntar “Quem protege o interesse público no mundo digital?” (BRITISH LIBRARY, [s/d]). As referidas restrições de acesso se manifestam em formas de *copyright*, direitos autorais e direitos de reuso, uma tríade que constitui conjuntamente o tema central do presente artigo.

Com a gradual consolidação do acesso aberto, temos hoje o que Willinsky (2006, p. 42) denomina um “domínio público de ciência em duas camadas” (“*a two-tiered public domain of science*”): uma baseada em acesso restrito pelo pagamento de taxas de acesso e a outra baseada em acesso aberto, com uma minoria de artigos (apenas em torno de 25% do total publicado anualmente) ocupando espaço simultaneamente nas duas camadas, graças àqueles pesquisadores-autores de artigos científicos publicados que também depositam uma cópia do mesmo num repositório digital (institucional ou disciplinar) em acesso aberto. Este cenário de “auto-arquivamento” em repositórios digitais de fato descreve o que é conhecido como a “via verde” de acesso aberto, que se compara com a outra, que é a “via dourada”, que consiste no conjunto grande e crescente de revistas científicas publicadas e disponibilizadas em acesso aberto. Esta segunda via é assim constituída pela criação e gestão de revistas



científicas que “nascem digitais” e, também, em termos da sua disseminação, nascem como publicações acesso aberto.

Embora ainda haja muito a ser feito para tornar acessível livre e abertamente muito da produção científica hoje existente no Brasil, já é fato que o acesso aberto e seus princípios estão se propagando e se consolidando no país, especialmente nas Instituições de Ensino Superior (IES). Várias destas tem instituído seu Repositório Institucional (RI), às vezes denominado como “Biblioteca Digital”, como é o caso da USP, com sua “Biblioteca Digital de Produção Intelectual – BDPI”. O *OpenDOAR* – um diretório internacional de repositórios digitais em acesso aberto, mantido pelo projeto SHERPA da Universidade de Nottingham no Reino Unido – atualmente registra 85 repositórios de acesso aberto no Brasil. Mas é inegável que comparativamente, a via dourada de acesso aberto seja bem mais consolidada no Brasil, que tem uma quantidade formidável de revistas científicas hoje disponíveis em acesso aberto, e ocupa o segundo lugar no mundo com 937 revistas registradas no *Directory of Open Access Journals* (DOAJ), somente ultrapassado pelos Estados Unidos que ocupam o primeiro lugar no mesmo Diretório, com 1218 revistas acesso aberto (DOAJ, 2014)ⁱ. Podemos afirmar que o grande responsável para este fortalecimento da via dourada no Brasil, comparada com a via verde ainda em estado incipiente, é o SciELO – *Scientific Electronic Library Online*.

O SciELO foi lançada em 1997 como um programa especial da FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), em parceria com a BIREME (Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde), sendo que o CNPq também passou a apoiar a SciELO em 2002. Vários outros países da América Latina e do Caribe integram a rede SciELO, além da África do Sul, Espanha e Portugal (PACKER, 2009). Como o próprio idealizador da SciELO, Abel Packer observa, a motivação principal do SciELO é de dar visibilidade mundial à “ciência perdida” e muito valiosa, que origina nos países em desenvolvimento, que por vários motivos, sempre sofreram maiores obstáculos para publicar nos canais das revistas científicas prestigiosas dos países centrais. A SciELO não se reduz a uma plataforma, repositório ou portal de acesso às revistas nela contidas: consiste também numa concepção e metodologia de editoração e publicação científica que garantem a qualidade das publicações disponibilizadas na SciELO. Tais padrões altos de qualidade tem acarretado numa situação louvável hoje em que vários dos índices bibliográficos internacionais e regionais (incluindo o *Web of Science*, Scopus, CROSSREF, *Google Scholar*, PubMed, LILACS, AGRIS, DOAJ) mantêm *links* atualizados para as revistas do SciELO, o que tem contribuído para o crescimento do Fator de Impacto de vários títulos do SciELO (PACKER, 2009). Em julho de 2012, o SciELO foi integrado no *Web of Knowledge* de Thomson Reuters. Vale ressaltar que desde a sua inepção, o programa de SciELO priorizou colocar as revistas online, e, como Packer (2009) observa, logo perceberam que o caminho de acesso aberto foi o caminho certo a seguir.



Vale ressaltar que fora da América Latina, existem editoras extremamente respeitadas que publicam somente em acesso aberto, a *Public Library of Science* (PLoS) e *Biomed Central* sendo exemplos que publicam artigos originais em acesso aberto principalmente na literatura STM (*Science, Technology & Medicine*). Os desafios para as editoras em acesso aberto se referem ao modelo de sustentabilidade financeira que utilizam, e o de estabelecer uma boa reputação para influir no fator de impacto do título da revista, ou mais comumente hoje, nas “*altmetrics*” dos *downloads* de artigos individuais. Os autores procuram veículos de publicação que otimizem a disseminação do seu trabalho na comunidade de pares, para que possa ser debatido, avaliado, e usado como base para futuros avanços, além de construir sua reputação dentro da comunidade científica. Existem vários levantamentos que demonstram que falta a consciência dentro da comunidade científica de que tais metas podem, sim, ser alcançadas via a publicação acesso aberto. Infelizmente, os modelos de negócios em acesso aberto conhecidos como “o autor paga” (*author pays open access*) tem sido igualados, na mente dos pesquisadores, com qualidade inferior. Cabe mencionar aqui que muitas das editoras comerciais tradicionais, que operam predominantemente sob o modelo de assinatura, também oferecem a chamada “opção aberta” de publicação: ou seja, no mesmo volume e número de uma determinada revista, encontra-se artigos em acesso aberto ao lado de artigos “fechados”, simplesmente porque o autor (ou seu órgão de fomento à pesquisa, ou a sua instituição empregadora) pagou para que o artigo pudesse ser disponível em acesso aberto imediatamente após publicação. As grandes editoras comerciais também têm seus títulos abertos (p.ex. *Springer Open Choice Publishing, Nature Open*). Morrison (2009, p. 119-120) enfatiza que é importante desmistificar a noção de que *toda* publicação acesso aberto cobra uma taxa para a publicação: cita um estudo de Hooker, de 2007, que demonstrou que 67% das revistas listadas no DOAJ não cobram pela publicação, e cita outro estudo de Suber e Sutton do mesmo ano, que demonstra que 83% das editoras acesso aberto de sociedades e associações científicas não cobram taxas de processamento de artigos.

Suber (2012) discorre sobre as características das revistas acesso aberto, enfatizando que a maioria segue as políticas editoriais e práticas das revistas convencionais, e de fato, o rigor do processo de editoração e publicação de revistas acesso aberto não deveria ser duvidado, como os padrões altos da SciELO atestam. As revistas acesso aberto têm a diferença evidente de que são livremente acessíveis para o leitor. Mas ser “aberto” significa apenas isso: ser livremente acessível *online*? Os idealizadores e mantenedores do hoje muito prestigioso Diretório de Revistas em Acesso Aberto (DOAJ – *Directory of Open Access Journals*), lança mão da definição seminal e “forte” de acesso aberto, consagrada na Declaração da *Budapest Open Access Initiative* (BOAI) de 2002, afirmando que:



Da definição BOAI de “acesso aberto”, apoiamos os direitos dos usuários a “ler, baixar, copiar, distribuir, imprimir, realizar buscas em, ou *linkar* aos textos completos desse artigos” como mandatórios/obrigatórios para que uma revista seja incluída no Diretório (DOAJ, 2014, tradução nossa)¹.

Ao que parece, essa definição “clássica” e “holística” de acesso aberto não tem sido unanimemente abraçado por editores, autores e pesquisadores. Existe um espectro variado de permissões, declarações de *copyright*, direitos autorais etc. vinculado às publicações disponíveis em plataformas como a SciELO e, inclusive, cadastradas no DOAJ. A questão no que concerne o grau de “abertura” é discutivelmente tão nebulosa que um conglomerado de organizações ativas em promover a publicação científica em acesso aberto – o SPARC, PLoS e OASPA – recentemente publicou um folheto intitulado *HowOpenIsIt?*² (“QuãoAbertoÉ?”) que esquematiza sucintamente atributos e critérios que podem ser usados para julgar se um dado documento ou publicação pode ser caracterizado como “acesso aberto” de fato. O intuito do panfleto é de ir além da ideia de que o acesso aberto se refere apenas à “gratuidade de acesso para todos os leitores” que tem acesso à internet; quer recuperar e revitalizar as dimensões do acesso aberto como definido originalmente na BOAI, ou seja, aquelas que também se referem aos direitos de uso, e reuso, além de esclarecer para os autores-pesquisadores quais são seus direitos se a publicação for realmente classificada como sendo de “acesso aberto”.

Diante do colocado, o estudo empírico realizado em escala modesta aqui apresentado foi motivado pela questão “Quão abertas são as revistas acesso aberto da SciELO do ponto de vista dos direitos autorais, *copyright* e dos direitos de reuso?” O intuito de tal enfoque é que os direitos que o autor detém de um artigo seu determinarão o que esse autor poderá fazer com seu artigo após publicado numa revista acesso aberto. Ou seja, quais são os direitos de reuso do artigo da sua autoria, que incluirão o seu direito de auto-arquivar uma cópia num repositório digital, por exemplo – nada mais do que a passagem da “via dourada” para a “via verde”. Como veremos, o fato de ter publicado numa revista acesso aberto não implica que essa passagem seja automática e livre de questionamentos. Assim, os objetivos do estudo foram os de:

- a) revisar uma parte da literatura sobre acesso aberto no que diz respeito aos direitos autorais, *copyright* e reuso;
- b) levantar, a partir de uma pequena amostra de revistas científicas disponíveis no SciELO-Brasil, as políticas de direitos autorais das mesmas;
- c) analisar essas políticas para averiguar o conceito de acesso aberto explícita ou implicitamente usado nas revistas levantadas, sinalizando, de certa forma, o grau de “abertura” destas.

¹ “From the BOAI definition of “open access”, we support the rights of users to “read, download, copy, distribute, print, search, or link to the full texts of these articles” as mandatory for a journal to be included in the directory.” Disponível no site do DOAJ em <http://doaj.org/about>.

² *HowOpenIsIt?* – disponível em: <http://www.sparc.arl.org/resource/howopenisit>



Posto isso, as seções a seguir apresentarão os passos seguidos em nosso levantamento.

MATERIAIS E MÉTODOS

A metodologia predominante neste estudo foi a de pesquisa documental, baseada em leitura de documentos e páginas da internet relevantes ao estudo. Para a coleta de dados deste estudo para averiguar as políticas autorais e de uso, e o grau de abertura das revistas acesso aberto, tiramos nossa amostra da plataforma SciELO-Brasil. A amostragem envolveu o levantamento de títulos de periódicos correntes das áreas de Ciências da Saúde e Ciências Biológicas, o que correspondeu a um total de 117 títulos de periódicos correntes. Foram eliminados os quatro títulos duplicados entre as duas listas antes de chegar a este total.

Após este levantamento, conferiu-se quais dos periódicos levantados do SciELO estão indexados no DOAJ, usando a ferramenta de busca no site do DOAJ que permite buscar apenas títulos de revistas científicas do mundo todo consideradas como sendo de acesso aberto.

Além disso, foi checado, na página da capa de cada revista (ou sob a opção de menu nesta, “sobre nós”) a partir da plataforma SciELO, se o conteúdo da revista é licenciado sob uma licença aberta ou não, já que no DOAJ esta informação nem sempre constava no registro da revista.

A última checagem foi na seção da revista chamada “Instruções aos autores”. Foram utilizados os termos de busca nesse texto, às vezes extenso, “*copyright*”, “direitos”, “autorais”, “cessão”, “termo”, “reservar”, “reservados”, e seus equivalentes no inglês quando as instruções para os autores foram dadas nessa língua.

Todos os dados levantados foram registrados numa planilha (do software Microsoft Excel) para sua padronização e subsequente manipulação e análise.

RESULTADOS

A seguir, apresentamos alguns dos resultados quantitativos levantados, com as ponderações e interpretações iniciais feitas em relação ao objetivo principal determinado ao início da pesquisa, que foi o de identificar o grau de “abertura” de revistas científicas acesso aberto hospedadas no SciELO-Brasil. A Tabela 1 abaixo apresenta os dados relacionados à amostra trabalhada.

Tabela 1 Detalhes da amostra das revistas

	Total de Títulos Correntes no SciELO	Número de títulos registrados no DOAJ	Número de títulos que <i>não</i> constam no DOAJ
Ciências da Saúde	91	83	8
Ciências Biológicas	26	23	3
	117	106	11

Focando exclusivamente nos 106 títulos que estão cadastrados no DOAJ – pois o fato de estarem indexados nesse Diretório significa que os avaliadores deste entendem que satisfazem os critérios para serem considerados como revistas acesso aberto – procedemos com um levantamento da Política Editorial das revistas, especificamente analisando o quesito de Propriedade Intelectual do título em questão. Sem exceção, *todos* os títulos examinados a partir da sua página de capa na plataforma SciELO-Brasil, tanto a partir da opção de menu “Sobre nós” quanto sob a seção “Propriedade Intelectual”, explicitam que: “Todo o conteúdo do periódico, exceto onde está identificado, está licenciado sob uma Licença *Creative Commons* do tipo atribuição BY-NC.” Entende-se que esta redação deve ser uma “frase padrão” utilizada pelo SciELO para todos os títulos de periódicos hospedados na sua plataforma, a qual comunica aos seus leitores e usuários a sua adesão aos princípios de acesso aberto em relação ao uso do material encontrado no SciELO.

É relevante lembrar aqui que a organização global sem fins lucrativos, a *Creative Commons* (CC), tem como objetivo deixar os autores e criadores determinarem sob quais termos e condições suas obras podem ser utilizados e quais direitos querem ceder e sob quais condições o fariam na forma de um espectro de licenças abertas das quais os criadores podem escolher. Como Silva e Lara (2011) observam, em contraste com o sistema tradicional legal de proteção dos direitos autorais, cujo dogma pode ser encapsulado na frase “todos os direitos reservados”, as licenças CC reconhecem que nem todo autor quer reservar todos os direitos para si, ou seja, é uma situação de “alguns direitos reservados”, e que isso é explicitado na licença CC de antemão:

Si en el sistema del derecho de autor el principio es que toda utilización de una obra debe tener un permiso expreso del titular de los derechos de autor, em el caso de las licencias CC el principio es que dicho permiso há sido entregado de antemano por el titular (SILVA; LARA, 2011, p. 18).

Não há espaço aqui para versar sobre todos os tipos de licenças CC³, mas cabe dizer que a licença CC BY-NC dos periódicos SciELO permite que:

³ Vide <http://creativecommons.org.br/as-licencas/>

outros *remixem*, adaptem, e criem obras derivadas sobre a obra licenciada, sendo vedado o uso com fins comerciais. As novas obras devem conter menção ao autor nos créditos e também não podem ser usadas com fins comerciais, porém as obras derivadas não precisam ser licenciadas sob os mesmos termos desta licença (CREATIVE COMMONS BRASIL, 2014).

Desse modo, a opção por uma licença CC pelo criador, autor ou instituição responsável pela publicação, comunica aos usuários do conteúdo ao qual a licença aplica, qual uso poderá, legitimamente, fazer desse conteúdo. Como se percebe da definição da licença CC BY-NC acima, é uma das licenças bastante flexíveis e abertas, sendo que a licença BY seria a mais flexível, exigindo apenas a atribuição / reconhecimento da autoria da obra, mas deixando o usuário bastante brecha para reutilizar o trabalho.

No entanto, quando jogamos luz no lado do autor, no quesito dos direitos autorais, a situação parece ser bastante distinta. Pois, mesmo que, como já dito, todos os títulos aqui analisados no SciELO-Brasil sejam portadores das licenças CC BY-NC, ao submeter o artigo ao periódico, na maior parte das vezes, os próprios autores têm que assinar um termo de cessão dos direitos autorais. A Tabela 2 abaixo apresenta os dados levantados no pequeno estudo aqui apresentado:

Tabela 2 Periódicos no SciELO que exigem cessão de direitos autorais

	Total de Títulos Correntes no SciELO indexados no DOAJ	Licença CC BY-NC no SciELO	Total de Títulos que exigem a cessão dos direitos autorais ao submeter o artigo	Total de títulos que <u>não</u> exigem a cessão dos direitos autorais	Não consta informação sobre direitos autorais de forma explícita
Ciências da Saúde	83	83	65 (78%)	4	14
Ciências Biológicas	23	23	9 (39%)	4	10
	106	106	74	8	21

A Tabela 2 mostra que 78% dos 83 revistas da área de Ciências da Saúde cadastradas no DOAJ – diretório este que, lembremos, adere-se aos princípios mais “fortes” de acesso aberto – exigem que os autores dos artigos assinem um termo de cessão dos direitos autorais. Esta porcentagem é de 34% das 26 revistas no DOAJ na área das Ciências Biológicas. Ao ceder os direitos autorais à revista científica, o(s) autor(es) basicamente transfere à revista, ou à instituição responsável pela revista, direitos exclusivos do direito patrimonial que a lei confere ao titular (autor): isto é, os direitos patrimoniais são transferíveis (diferentemente dos direitos morais). Tais direitos patrimoniais incluem, entre outros, o direito de distribuição, adaptação, tradução, a disseminação pública (por exemplo, na internet), ou publicação (SILVA; LARA, 2011).

Engloba também, os direitos de posterior utilização pelo autor, incluindo o depósito num repositório digital, por exemplo (MELERO, 2010).

Os termos de cessão dos direitos autorais empregados por estas revistas em acesso aberto do SciELO e registrado no DOAJ englobam desde a “mera” transferência de *copyright* do(s) autor(es) para a revista, às vezes com a revista permitindo com isso a “reprodução do todo ou em parte desde que mencionadas as fontes”, até a mais abrangente e conservador necessidade de transferir todos os direitos, que passam a ser “propriedade exclusiva” da revista, vedando a reprodução parcial ou total sem a devida autorização desta. Existe, inclusive, o caso de um título em saúde que gerencia e veicula a revista a partir da Plataforma do *Open Journal System (OJS)* do *Public Knowledge Project* (e indexada e acessível a partir do SciELO), o conteúdo do qual está licenciado sob a licença bastante aberta, a CC BY 3.0⁴, mas cujo termo de cessão de direitos autorais é bastante rígido, como visto no trecho a seguir (no qual o “cedente” é o autor, e a “cessionária”, a revista:

3. O CEDENTE cede e transfere todos os direitos autorais relativos à OBRA à CESSIONÁRIA, especialmente os direitos de edição, de publicação, de tradução para outro idioma e de reprodução por qualquer processo ou técnica. A CESSIONÁRIA passa a ser proprietária exclusiva dos direitos referentes à OBRA, sendo vedada qualquer reprodução, total ou parcial, em qualquer outro meio de divulgação, impresso ou eletrônico, sem que haja prévia autorização escrita por parte da CESSIONÁRIA (VISA EM DEBATE, 2014).

Uma certa contradição que existe com a aplicação de termos de cessão de direitos deste teor, ao lado de licenças abertas como as do CC, é que do ponto de vista do usuário, pressupõe-se que as permissões de uso são flexíveis mas que para os autores, as permissões em relação àquilo que pode fazer com seu próprio texto são bem mais restritos.

A última coluna da Tabela 2 acima (na página anterior) contem dados referentes àquelas políticas editoriais que não explicitam, no corpo do texto da política, quais direitos o autor pode continuar detendo após ter seu artigo publicado no periódico. Evidentemente, o fato de não explicitar essa parte da política não implica que o autor se manterá como detentor dos seus direitos autorais. Há de se observar que justamente nesses títulos de periódicos – como é o caso na maioria das revistas científicas online hoje – o processo de submissão do manuscrito é inteiramente automatizado e consiste em vários passos em que o autor fornece informações e também, às vezes, tem que aceitar certos termos da revista em questão, “assinalando caixinhas”. E é comum que nesses passos automatizados, o autor terá que aceitar a cessão dos direitos autorais à revista. Existe a possibilidade, então, que estas 21 revistas a mais requeiram a cessão dos direitos autorais, mas infelizmente, não houve tempo neste estudo para realizar testes para comprovar ou refutar essa hipótese.

⁴ http://creativecommons.org/licenses/by/3.0/deed.pt_BR

Entre as políticas editoriais das nove revistas das duas áreas que explicitamente não exigem a cessão dos direitos autorais à revista, no geral, podemos afirmar que representam políticas editoriais de revistas científicas que entenderam de modo mais “holístico” e coerente as implicações de ser uma revista acesso aberto, que é hospedada, divulgada e acessada a partir do SciELO e, entre outros índices, indexada no DOAJ. Podemos exemplificar estas, na área das Ciências Biológicas, com as revistas *Journal of Venomous Animals and Toxins including Tropical Diseases* e *Ambiente & Sociedade*. A primeira claramente declara que “O copyright fica com os autores”, e remete o autor ao site da Editora (*Publisher*) em acesso aberto, o BioMed Central, para maiores detalhes, onde está claramente explicado (em inglês) o quê o autor e usuário do artigo podem fazer com o artigo⁵.

O segundo caso, da revista *Ambiente & Sociedade*, é um que podemos afirmar como sendo a revista mais aberta encontrada entre os títulos das duas áreas analisadas. Vale até a pena citar o trecho na íntegra pois encapsula algumas qualidades ideais de uma política editorial, posto que ao mesmo tempo em que informa aos autores sobre o que implica aceitar uma licença CC, elucida também alguns princípios e vantagens de acesso aberto:

Autores que publicam neste periódico concordam com os seguintes termos:

1. Autores mantêm os direitos autorais e concedem ao periódico o direito de primeira publicação, com o trabalho simultaneamente licenciado sob a Licença Creative Commons Attribution que permite o compartilhamento do trabalho com reconhecimento da autoria e publicação inicial neste periódico.
2. Autores têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada neste periódico (ex.: publicar em repositório institucional ou como capítulo de livro), com reconhecimento de autoria e publicação inicial neste periódico.
3. Autores têm permissão e são estimulados a publicar e distribuir seu trabalho online (ex.: em repositórios institucionais ou na sua página pessoal) a qualquer ponto antes ou durante o processo editorial, já que isso pode gerar alterações produtivas, bem como aumentar o impacto e a citação do trabalho publicado (Veja O Efeito do Acesso Livre) (AMBIENTE & SOCIEDADE, 2014).

Relativo ao espectro do grau de abertura de revistas e textos acesso aberto apresentado no supramencionado panfleto *HowOpenIsIt?*, a *Ambiente & Sociedade* se localiza claramente do topo do espectro, de ser clara e explicitamente aberta na sua política editorial, deixando claro para os usuários e autores o que são livres a fazer com o seu artigo pós-publicação.

⁵ BioMed Central license agreement. Disponível em: <http://www.biomedcentral.com/about/license/>



Entre as também poucas (4) revistas na área de Ciências de Saúde que não explicitamente mencionam a necessidade do(s) autor(es) cederem os direitos autorais, podemos citar a *Brazilian Dental Journal*. Na página da política editorial e instruções aos autores, em nenhum momento há menção feita de *copyright* ou direitos autorais, e a submissão do manuscrito não é automatizada, mas pelo envio “manual” para um e-mail. Na página de capa da revista, mostra quem são os órgãos patrocinadores e “mantenedores” da revista, que incluem a CNPq, CAPES, o MEC e MCT; então surpreender-se-ia se a revista operasse sob um regime mais fechado e conservador de direitos autorais, já que esses órgãos do governo brasileiro estimulam a adoção de iniciativas que fomentem a abertura da ciência brasileira, sobretudo na área de saúde.

CONSIDERAÇÕES PARCIAIS

O estudo modesto aqui apresentado constituiu, para as autoras, uma forma de testar uma suposição intuída a partir de conversas informais e entrevistas realizadas com pesquisadores científicas durante outra pesquisa realizada sobre o acesso aberto, a saber: que ainda existe desconhecimento e confusão entre eles em torno do termo “acesso aberto”, confusão esta que pode ser reforçada pela ambiguidade das políticas editoriais de revistas acesso aberto. Entendemos que estas são considerações parciais, pois gostaríamos de realizar futuramente outra etapa da pesquisa, consistindo em entrevistar os editores de uma amostragem das revistas analisadas, além de analisar mais títulos de outras áreas no SciELO-Brasil.

A contradição em torno dos direitos autorais e direitos de uso das revistas aqui examinadas reside principalmente no fato de que ao mesmo tempo em que todas as revistas, sem exceção, operam sob uma licença Creative Commons BY-NC, a maioria delas também exige que os autores dos artigos transfiram os direitos autorais às revistas. Há de esclarecer aqui que não existe uma regra que dite que revistas acesso aberto não podem ser os detentores de tais direitos só por ser acesso aberto, mas também há necessidade de esclarecer que quando um *usuário* (leitor) dessas revistas vê que o título carrega uma licença CC BY ou CC BY-NC, entenderá que ele mesmo tem a liberdade de fazer certas coisas com este texto que a licença estipula. É importante reconhecer que o usuário-leitor não tem a obrigação de procurar saber se o título em questão obrigou o *autor* ceder seus direitos, sobretudo se estas obrigações se encontram entre as “instruções aos autores” e são dirigidas ao autor no momento da submissão do artigo. No entanto, há de se reconhecer que o tom da redação utilizada em muitos desses termos de concessão dos direitos autorais é bastante rígida e legalista, não deixando muita margem de interpretação equivocada para os autores que os aceitem. Mas voltando à indagação: e se o usuário-leitor apenas ver que o título é licenciado sob uma licença flexível de CC?



De acordo com o espectro de “abertura” de documentos e revistas apresentado no panfleto *HowOpenIsIt?* (“Quão aberto é?”) lançado em 2011 por SPARC, PLoS e OASPA, as revistas acesso aberto aqui analisadas se enquadram na parte do espectro mais aberto ao falar dos direitos de reuso pelos usuários, mas ao mesmo tempo, a maioria delas está enquadrada apenas um passo do ponto no espectro que marca a posição mais fechada e conservador no que diz respeito aos direitos autorais, pela qual a editora (*Publisher*) detém o *copyright*, às vezes com algumas permissões pelo reuso da versão publicada pelo autor.

Em suma, pode ser concluída que ainda existe necessidade por maiores esclarecimentos com editores de revistas acesso aberto sobre o que, exatamente, esta modalidade de publicação implica, e deveria implicar, tanto para usuários, quanto para autores, para que as políticas editoriais de tais revistas possam ser cada vez mais coerentes e assim, trabalhar em prol da circulação e alcance de novas pesquisas científicas brasileiras.

REFERÊNCIAS

AMBIENTE & SOCIEDADE. Declaração de direito autoral. Disponível em: <http://submission.scielo.br/index.php/asoc/about/submissions#onlineSubmissions>; acesso em: 8 ago 2014.

BRITISH LIBRARY. **Driving UK research. Is *copyright* a help or a hindrance: a perspective from the research community.** British Library. London. s/d.

CREATIVE COMMONS. Disponível em: <http://creativecommons.org.br/o-que-e-o-cc/>. Acesso em 2 ago.2014.

DOAJ. DIRECTORY OF OPEN ACCESS JOURNALS. Disponível em: <http://doaj.org/oainfo>

GUIBAULT, L. Owning the right to open up access to scientific publications. In: GUIBAULT, L.; ANGELOPOULOS, C. **Open content licensing: from theory to practice.** Amsterdam: Amsterdam University Press, 2011. p. 137-168.

HARNAD, S. Opening access by overcoming Zeno’s paralysis. In: JACOBS, N. **Open Access: key strategic, technical and economic aspects.** Oxford: Chandos, 2006. p. 73-86.



KNEZO, G. **Open access publishing and citation archives: background and controversy**. Congressional Research Service. Washington D.C., p. 51. 2006.

MELERO, R. **Guia práctica sobre los derechos patrimoniales o de explotación (*copyright*) y su relación con el auto-archivo en repositorios de acceso abierto**. [S.l.]: [s.n.], 2010. Disponível em: <<http://www.accesoabierto.net/es/bibliografia/guia-practica-sobre-los-derechos-patrimoniales-o-de-explotacion-copyright-y-su-relacion>>. Acesso em: 4 ago 2014.

MONBIOT, G. The Lairds of Learning. **The Guardian** 30th August 2011. Disponível em: <http://www.monbiot.com/2011/08/29/the-lairds-of-learning/> ; acesso em: 1 ago 2014.

MORRISON, H. **Scholarly Communication for Librarians**. Oxford: Chandos, 2009.
PACKER, A. The SciELO Open Access: A gold way from the South. **Canadian Journal of Higher Education**, v. 39, n. 3, p. 111-126, 2009.

PINFIELD, S. Paying for open access? Institutional funding streams and OA publication charges. **Learned Publishing**, v. 23, n. 1, p. 39-52, 2010.

RODRIGUES, E. Acesso livre ao conhecimento: a mudança do sistema de comunicação da ciência e os profissionais de informação. **Cadernos BAD**, n. n.1, 2004. p. 24-35.

SILVA, A. C.; LARA, C. **Políticas Editoriales de Publicaciones Académicas en Línea en Latinoamérica**. Santiago : Derechos Digitales, 2011. Disponível em: <https://www.derechosdigitales.org/que-hacemos/publicaciones/> ; acesso em: 4 ago 2014.

SUBER, P. **Open Access**. Cambridge, MA: MIT Press, 2012.

VIGILÂNCIA SANITÁRIA EM DEBATE. **Declaração de Direito Autoral**. Disponível em: <http://www.visaemdebate.incqs.fiocruz.br/index.php/visaemdebate/about/submissions#copyrightNotice>; acesso em: 15 jul 2014.

ⁱ <http://doaj.org/>



A PRESERVAÇÃO DO PATRIMÔNIO ARQUIVÍSTICO FOTOGRÁFICO E O ACESSO ABERTO UTILIZANDO O SOFTWARE DE DESCRIÇÃO ICA-AtoM

THE ARCHIVAL PHOTOGRAPHIC PRESERVATION OF HERITAGE AND THE OPEN ACCESS USING SOFTWARE DESCRIPTION ICA-ATOM

Henrique Machado dos Santos¹, Dhion Carlos Hedlund², Daniel Flores³

¹Acadêmico do Curso de Arquivologia Universidade Federal de Santa Maria – UFSM

henrique.gralha@gmail.com

²Me, Professor do Curso de Arquivologia Fundação Universidade Federal do Rio Grande (FURG)

dhion.hedlund@gmail.com

³Dr, Professor do Curso de Arquivologia Universidade Federal de Santa Maria – UFSM

dfloresbr@gmail.com

RESUMO

A Arquivologia vem ganhando força com o surgimento de novas tecnologias da informação e comunicação, isso faz o arquivista ser visto como um administrador de informações, mudando assim aquela imagem antiquada. Funções do profissional como a descrição, a preservação e a garantia de acesso à informação, podem ser realizadas com o auxílio de softwares específicos. Nesse contexto o Curso de Arquivologia da Universidade Federal de Santa Maria (UFSM) juntamente ao Departamento de Arquivo Geral da UFSM, desenvolveram práticas arquivísticas voltadas à instituição. Foi instalado o ICA-AtoM, um *software* livre, o qual possui os requisitos desejados para atender as necessidades de descrição e difusão do acervo fotográfico existente. Essa ideia resultou em um projeto de extensão, que conta com bons resultados. Um dos fatores diferenciais do trabalho é que o ICA-AtoM é um software específico para a atividade de descrição arquivística, ou seja, não há adaptações em seu uso.

Palavras-Chave: Difusão. Preservação. Tecnologia da informação. Descrição. Fotografia.



ABSTRACT

The Archival been gaining momentum with the emergence of new information technologies and communication, it makes the archivist be seen as an administrator of information, thereby changing that outdated image. Professional functions such as description, preservation and ensuring access to information, can be performed with the aid of specific software. In this context the Course of Archival Science at the Federal University of Santa Maria (UFSM) along with the Department of General File UFSM developed aimed at institution archival practices. ICA-AtoM, free software, which has desired to meet the needs of description and dissemination of existing photographic collection requirements was installed. This idea resulted in an extension project, which has good results. One of the key differentiators of the work is that the ICA-Atom is a specific software for archival description of the activity, ie, there are no adjustments in their use.

Keywords: Dissemination. Preservation. Technologie of information. Description. Photographic.

INTRODUÇÃO

O Departamento de Arquivo Geral (DAG) da Universidade Federal de Santa Maria (UFSM) possui um acevo fotográfico que constitui um conjunto documental arquivístico. Esses documentos possuem um valor relevante com repercussão para a comunidade local, regional e até mesmo nacional.

Tendo em vista a necessidade de preservar e difundir esse patrimônio, foi realizada uma parceria entre o Curso de Arquivologia da UFSM e o DAG, resultando em um projeto de extensão. O projeto “A Preservação do Patrimônio Cultural Regional de Santa Maria na Produção de Instrumentos de Pesquisa Arquivísticos” com a finalidade de digitalizar os negativos fotográficos do acervo existente no DAG, gerando um formato TIFF¹ para preservação, e uma derivada de acesso no formato JPEG².

Este artigo tem por objetivo apresentar os resultados da execução do projeto de extensão tendo em vista a implementação do *software* livre para descrição arquivística, o ICA-AtoM³. A partir destes resultados obtidos durante o projeto, pretende-se enumerar vantagens proporcionadas pelo ICA-AtoM para a Arquivologia.

¹ É um formato de arquivo raster para imagens digitais, que é usado extensamente em aplicações de manipulação de imagem e também muito utilizado para o intercâmbio de imagens entre as diversas plataformas.

² É um método comumente usado de compressão com perdas para imagens digitais, especialmente para aquelas imagens produzido pela fotografia digital .

³ *International Council on Archives – Access to Memory*



METODOLOGIA

Do ponto de vista metodológico esta pesquisa é de natureza aplicada, pois pode ser aplicada a uma situação do mundo real (SILVA & MENEZES, 2005). Sua abordagem é qualitativa, pois conforme Silva & Menezes (2005) a pesquisa qualitativa considera que:

[...] há uma relação dinâmica entre o mundo real e o sujeito, isto é, um vínculo indissociável entre o mundo objetivo e a subjetividade do sujeito que não pode ser traduzido em números. [...] Não requer o uso de métodos e técnicas estatísticas. O ambiente natural é a fonte direta para coleta de dados e o pesquisador é o instrumento-chave. É descritiva. [...] (SILVA & MENEZES, 2005, p.20).

Os resultados encontrados tem como base a prática experimental, e são relatados conforme as etapas realizadas durante a implementação do *software* ICA-AtoM. Assim partindo de uma breve revisão de literatura verificou-se a conformidade do ICA-AtoM com as normas para descrição arquivística.

A base de dados do ICA-AtoM foi hospedada em um servidor localizado no Departamento de Arquivo Geral, onde se utilizou o sistema operacional Linux Ubuntu 11.04, plataforma 64 bits, mantido no ar pelo CPD (Centro de Processamento de Dados) da UFSM.

Para o funcionamento do ICA-AtoM, foi necessária a adesão de determinados softwares: o servidor Apache⁴, a base de dados MySQL⁵, a linguagem de programação PHP⁶, o Symfony⁷ e o Qubit, este usado inicialmente para desenvolver o ICA-AtoM. Nenhum dos *softwares* que foram citados possui licenças comerciais ou que possuam algum direito de propriedade pagas, isso faz com que o uso do ICA-AtoM seja totalmente sem custos.

⁴ Servidor web livre

⁵ É um sistema de gerenciamento de banco de dados (SGBD), que utiliza a linguagem SQL (Linguagem de Consulta Estruturada).

⁶ É uma linguagem interpretada livre, usada originalmente apenas para o desenvolvimento de aplicações presentes e atuantes no lado do servidor, capazes de gerar conteúdo dinâmico na World Wide Web.

⁷ Symfony é um framework web escrito em PHP.



RESULTADOS PARCIAIS

Neste artigo são relatadas as experiências vivenciadas com a implementação do ICA-AtoM, onde é evidenciada a sua conformidade com as normas para descrição arquivística. Paralelamente a este estudo foi analisada a parte técnico-operacional do software, procurando vantagens para aplicações em âmbitos institucionais. Outro aspecto a ser ressaltado é a tradução do ICA-AtoM versão 1.3 e a tradução do manual do usuário ICA-AtoM.

A necessidade de preservar, descrever e difundir o acervo

O Patrimônio Fotográfico do DAG encontra-se disponível ao público, normalmente sendo consultado por pesquisadores em geral da área de história como alunos e professores. Sendo assim, para a tarefa do arquivista de garantir o acesso às informações e também preservá-las, são geradas as imagens no formato TIFF com as respectivas configurações que são recomendadas para preservação de fotografias em meio digital, conservando assim o suporte de filmes negativos. Os filmes se encontram acondicionados em envelopes que estão separados por assuntos, cabendo aos bolsistas selecionar as fotos a serem digitalizadas de acordo com as respectivas recomendações de relevância do conteúdo e nitidez da imagem.

Inicialmente foi proposta uma base de dados modelada no software Microsoft Access, onde as fotos em formatos JPEG eram descritas. Claro que esse processo de busca diretamente nos negativos não deixa de existir, até porque como citado, nem todas as fotografias foram digitalizadas. No que se refere a base de dados em Access, é de boa estima ressaltar que sua construção foi fundamentada em campos obrigatórios da ISAD(G), sendo assim uma ferramenta de aplicabilidade satisfatória, a qual foi desenvolvida por bolsistas iniciantes do projeto.

A ideia inicial era desenvolver uma ferramenta própria para executar a descrição arquivística e a difusão na *internet*, sendo que não se obteve êxito no desenvolvimento dessa ferramenta. Então o ICA-AtoM começou a ser usado inicialmente em caráter de teste e mesmo assim a descrição continuou sendo realizada na base em Microsoft Access. Esta decisão foi tomada para evitar perdas indesejáveis da descrição caso o software ICA-AtoM não atingisse as expectativas.

A difusão do acervo fotográfico é realizada no local através de consultas e também via *internet*, onde é utilizado o software ICA-AtoM. Qualquer pessoa tem direito de solicitar uma pesquisa sobre um determinado fato ao Departamento, competindo ao mesmo verificar a existência de informações relevantes sobre o assunto no acervo de negativos. Esse é o principal tipo de busca, pois apenas parte do acervo está disponível na grande rede, pois nem todos os negativos fotográficos foram digitalizados, apenas as fotos mais relevantes como já citado. Em um conjunto de trinta fotografias, podemos escolher oito, por exemplo, de acordo com as regras já antes estabelecidas.

Foi escolhido esse método com a ideia de que a base de fotografias inserida no ICA-AtoM incentive as pesquisas ao acervo. De maneira geral, a base de dados do ICA-AtoM vem para servir como um mecanismo de amostragem do que poderá ser encontrado no Acervo do DAG, através da descrição das derivadas de acesso e difusão dessas informações via *internet*.

Cabe ressaltar que no software ICA-AtoM não foram armazenados formatos de preservação TIFF, devido ao tamanho do arquivo gerado.

O ICA-AtoM e as normas de descrição

O ICA-AtoM é um projeto elaborado pelo CIA – Conselho Internacional de Arquivos, juntamente a outros parceiros como a UNESCO⁸. A ideia surgiu devido à necessidade de ter a informática como ferramenta de trabalho no campo da arquivística, em virtude da disseminação do conhecimento e da busca constante por padronizações no campo das Ciências da Informação.

Segundo o manual online ICA-ATOM (2014), origem do termo vem de “*International Council on Archives – Access to Memory*” (Conselho Internacional de Arquivos – Acesso à Memória). Ele é um software livre que possui código aberto, que contempla as normas de descrições arquivísticas. O software é executado diretamente no navegador instalado, por exemplo: no Mozilla Firefox.

Principais características do ICA-AtoM:

- a) Possui um sistema multi-arquivos e multi-repositório, ou seja, poderemos ter diversos documentos de diversos fundos;
- b) É um software multilíngue, que permite a cada instituição fazer a sua própria tradução;
- c) É multi-plataforma, ou seja, está disponível para sistemas operacionais Linux e Windows, por exemplo;
- d) Contempla as normas internacionais para descrição arquivística: ISAD(G), ISAAR(CPF), ISDIAH e ISDF;
- e) Contempla também outras normas relacionadas como: EAD⁹, EAC, METS¹⁰, MODS¹¹ e Dublin Core¹²;
- f) Possui código aberto, possibilitando que qualquer pessoa com conhecimentos de programação possa manipular o código, contribuindo assim para o desenvolvimento do *software*;
- g) É um software livre, não há custos para usá-lo.

⁸ Organização das Nações Unidas para a Educação, a Ciência e a Cultura.

⁹ É uma norma para a codificação de instrumentos de descrição documental arquivísticos utilizando a linguagem XML mantida pela Biblioteca do Congresso e pela Society of American Archivists.

¹⁰ É um padrão de metadados para a codificação de metadados sobre objetos descritivos, administrativos e estruturais dentro de uma biblioteca digital, expressa por meio do XML linguagem de esquema do World Wide Web Consortium.

¹¹ É um XML baseada esquema de descrição bibliográfica desenvolvido pelo United States Library of Congress "Rede de Desenvolvimento e Normas Office.

¹² É um esquema de metadados que visa descrever objetos digitais, tais como, videos, sons, imagens, textos e sites na web. Aplicações de Dublin Core utilizam XML e o RDF.



O ICA-AtoM foi desenvolvido em torno das normas de descrição do International Council on Archives (ICA):

- a) General International Standard Archival Description (ISAD(G)) - 2nd edition, 1999;
- b) International Standard Archival Authority Record (Corporate bodies, Persons, Families) (ISAAR(CPF)) - 2nd edition, 2003;
- c) International Standard For Describing Institutions with Archival Holdings (ISDIAH) - 1st edition, March 2008;
- d) International Standard For Describing Functions (ISDF) - 1st edition, May 2007.

O CIA (2000, p.11) define a ISAD(G) como uma norma que “estabelece diretrizes gerais para a preparação de descrições arquivísticas. Deve ser usada em conjunção com as normas nacionais existentes ou como base para a sua criação”. Ou seja, a ISAD(G) vem para orientar o trabalho dos arquivistas, seja incorporando-se as normas existentes ou auxiliando no processo de criação das mesmas.

De acordo com o CIA (2003), a norma ISAAR(CPF) dá diretrizes para a preparação de registros de autoridade arquivística que forneçam descrições de entidades (entidades coletivas, pessoas e famílias) relacionadas à produção e manutenção de arquivos.

Com isso é possível descrever uma entidade coletiva, pessoa, ou família como unidades dentro de um sistema de descrição arquivístico, controlar a criação e uso de pontos de acesso em descrições arquivísticas e documentar relações entre diferentes produtores de documentos e entre essas entidades e os documentos que produziram e/ou outros recursos sobre ou produzidos por essas mesmas entidades. (CIA: ISAAR(CPF), 2003, p. 11).

A ISDF segundo o CIA (2007) é uma norma que dá diretrizes para a preparação de descrições de funções de entidades coletivas associadas à produção e manutenção de arquivos.

O termo “função” é usado nesta norma para incluir não somente funções, mas também qualquer uma das subdivisões de uma função, tais como subfunção, procedimento operacional, atividade, tarefa, transação ou outro termo de uso internacional, nacional ou local. A norma pode ser usada para descrever uma função ou qualquer de suas subdivisões. (CIA: ISDF, 2007, p. 11).

A ISDIAH é denominada pelo CIA (2008) como uma norma que apresenta regras gerais para a normalização de descrições de instituições com acervos arquivísticos. Permitindo assim:



O fornecimento de orientação prática na identificação e contato com instituições com acervos arquivísticos e no acesso ao acervo e aos serviços disponíveis; a elaboração de diretórios de instituições com acervo arquivístico e/ou listas de autoridade; o estabelecimento de conexões com listas de autoridade de bibliotecas e museus e/ou o desenvolvimento de diretórios comuns de instituições de patrimônio cultural nos níveis regional, nacional e internacional; e a produção de estatísticas de instituições com acervo arquivístico, nos níveis regional, nacional ou internacional. (CIA: ISDIAH, 2008, p. 11).

O ICA-AtoM possibilita efetuar as descrições arquivísticas, de acordo com o padrão internacional estipulado pelo CIA, sendo assim uma ferramenta que agrega eficiência e praticidade. Devido ao advento da tecnologia e a busca constante por melhorias e disseminação da informação, se faz necessário ao arquivista buscar essas ferramentas de trabalho e usá-las a seu favor.

O *software* ICA-AtoM foi programado conforme as normas internacionais pra descrição arquivística. Ele é ideal para o uso no campo arquivístico, por ser um *software* voltado exclusivamente para descrição arquivística, além de ser suficientemente flexível para a adaptação de outras normas de descrição. A partir desse conhecimento o ICA-AtoM ganha credibilidade entre os profissionais da área. Esses fatores impulsionaram os trabalhos realizados no decorrer do projeto, facilitando assim a descrição realizada nas fotografias digitalizadas que foram inseridas na base de dados do ICA-AtoM.

Vantagens do uso do ICA-AtoM na descrição

A aplicação do *software* traz consigo diversas facilidades referentes ao uso de uma ferramenta para descrição e na precisão do trabalho realizado. Devido exclusivamente as suas funcionalidades e a sua normalização. Observa-se que são mais que necessários os estudos das normas de descrição para operá-lo, porém com os estudos das normas é possível. Dentre as facilidades apresentadas pelo *software* é importante destacar:

- A possibilidade de descrição completa em todos os níveis de mais de um fundo documental;
- Abranger os campos das normas internacionais para descrição arquivística;
- Permitir a indexação de assuntos entre as descrições;
- Fazer *upload* de documentos digitais;
- Facilidade de difusão do acervo para a comunidade através da *internet*;
- Abrangência dos recursos presentes na base de dados que foi desenvolvida em Access, além da inclusão de novos recursos;
- Contas de usuários com permissões de acesso que são definidas pelo administrador;
- Sua interface é bem amigável, podendo ainda ser personalizada;

- Os conteúdos buscados por um pesquisador estão “linkados” sendo assim possível uma busca rápida e eficiente;
- O pesquisador pode usufruir de recursos como a pesquisa avançada que delimita os seus resultados;
- Não implica em possuir conhecimentos de programação para se uso e alimentação de conteúdo;
- Os usuários podem gerar relatórios das descrições e dos itens documentais;
- As descrições realizadas podem ser definidas como preliminares;
- Suporta importações e exportações XML;
- Permite marcar descrições como “preliminares”, uma propriedade que deixa oculto para os pesquisadores, é de boa ajuda em descrições incompletas ou que contenham incertezas.

De modo geral o ICA-Atom substituiu os modelos de descrição impressos, pela simples razão de possuir os campos pertencentes a ISAD(G) em sua estrutura. Além disso, o software conta com todos esses recursos já citados anteriormente, sendo assim um *software* que engloba a descrição arquivística e difusão de maneira eficaz. Estes processos se tornaram mais simples devido ao *software* possuir uma estrutura praticamente pronta para o uso. É possível descrever um conjunto de fotografias referentes a um determinado assunto, diretamente no *software*, sendo que este assunto especificado servirá como mecanismo de relação dos itens de um determinado tema.

Segundo Hedlund (2012) com o estudo do ICA-AtoM é possível implementar um sistema para descrição arquivística via web, para uma instituição ou acervo pessoal, atingindo milhões de pessoas em diversos idiomas, tendo um sistema seguro e de baixo custo, colocando em prática a funções arquivísticas de descrição e difusão do acervo.

Embora os termos do ICA-AtoM sejam baseados fielmente na ISAD(G) é válido sobressaltar que a aplicação da norma é realizada com base na NOBRADE. Que se apresenta da seguinte forma NOBRADE (2006) não é uma mera tradução da ISAD(G) e sim uma adaptação das normas internacionais para a realidade brasileira.

De acordo com Rezende, F. F. P & Fonseca, R. C. (2008): a norma ISAD(G) estabelece diretrizes gerais para preparação de descrições arquivísticas, devendo ser usada em conjunção com as normas nacionais existentes ou mesmo como base para sua criação.

Tradução dos termos do *software*

O ICA-AtoM é um software multilíngue, porém precisa que o usuário faça a tradução, pois inicialmente apenas o idioma em Língua Inglesa encontrava-se disponível. Posteriormente foram realizadas traduções para idiomas como o espanhol, o francês e recentemente para o português de Portugal. O Conselho Internacional de Arquivos é quem designa ao Arquivo Nacional de cada país, a possibilidade de efetuar a tradução oficial para cada língua respectivamente.

Atualmente o software se encontra na versão AtoM 2.0, mas a tradução dos termos teve início na versão 1.2, a qual não possuía tradução oficial para a língua portuguesa do Brasil. Por esse motivo se fez necessário às instituições brasileiras que usam o ICA-AtoM, realizar suas próprias traduções, como é o caso da Casa de Oswaldo Cruz, do Portal de Acervos Arquivísticos da Universidade Federal de Goiás (UFG) e do Patrimônio Arquivístico da UFSM. Posteriormente iniciou-se um processo de migração da versão 1.2 para a versão 1.3, neste caso a base de dados e os termos traduzidos até o momento foram preservados durante a migração.

O usuário que possuir permissões de acesso poderá traduzir os termos do software, porém não há um local específico para realizar a tradução, o usuário tem que navegar página por página traduzindo assim os termos referentes a cada uma. Se o usuário “logado” possuir permissões de tradutor, aparecerá na base da página uma barra amarela, onde clicando aparecerão os termos da respectiva página. Para traduzir basta digitar o novo termo no campo correspondente, é um procedimento simples, porém demorado.

No subprojeto realizado no DAG foram analisadas as traduções efetuadas por algumas páginas brasileiras, buscando fazer um comparativo. Também foram procurados termos na versão do português de Portugal, a fim de aproximar a clareza e a similaridade dos termos, até porque a ideia de unificação da língua já está presente em nossas recentes reformas ortográfica. Além das páginas citadas e da versão de Portugal, buscou-se o apoio das normas de descrição internacionais. A versão inicial do português de Portugal apresentava algumas falhas como falta de padrão entre os termos, foram encontrados na página referente a ISAD(G) termos como “Zona” e “Área”, que levam a um mesmo entendimento, além de possuir termos ainda não traduzidos.

Para finalizar o trabalho surge então uma atividade complementar: fazer a tradução do Manual do Usuário disponível no site do ICA-AtoM. Até o presente momento também não há uma tradução completa do manual do usuário, fato que dificulta o entendimento das funcionalidades do software. Já em relação as necessidades do DAG, seria muito mais prático ter o software e o manual em Língua Portuguesa, dessa forma facilitaria a aprendizagem dos futuros bolsistas do projeto.

A página da base de dados do Patrimônio Arquivístico da UFSM já está com os termos do ICA-AtoM (Versão 1.3) traduzidos para a Língua Portuguesa do Brasil. A tradução do Manual do Usuário foi concluída e esta disponível em <http://w3.ufsm.br/dag/images/ICA-AtoM%20manual%20do%20usuario%20-%20PT%20BR.pdf> (PAVEZI, 2013).

A partir do momento que software e manual estivessem traduzidos, o próprio ICA-AtoM seria mais difundido, obtendo maior aceitação entre as instituições brasileiras. Cabe ressaltar novamente que o Arquivo Nacional é o responsável pela tradução oficial, a versão obtida nesse projeto é para uso do próprio Patrimônio Arquivístico da UFSM.



CONSIDERAÇÕES PARCIAIS

O ICA-AtoM surge com um grande potencial na área da Arquivologia devido a incorporação das normas de descrição internacionais, conjuntamente a necessidade do profissional da área de informatizar os processos. Construído com a ideia de ser um software livre e de código aberto com necessidade apenas de auxílio de programas que não agregam custos e sendo assim voltado diretamente para o arquivista. Na época criava uma grande expectativa e hoje a sua popularização tornou-se realidade entre os profissionais da área.

O projeto: “A Preservação do Patrimônio Cultural Regional de Santa Maria na Produção de Instrumentos de Pesquisa Arquivísticos”, obteve bons resultados. Acessando a página do Patrimônio Arquivístico, UFSM (2014, Online) <www.ufsm.br/patrimonioarquivistico> o usuário poderá fazer pesquisas na base de dados, com a interface traduzida. Para maiores informações do projeto há também o blog do projeto, disponível no endereço virtual <patrimoniograficoufsm.blogspot.com.br>. O Manual do Usuário ICA-AtoM 1.3 em língua portuguesa do Brasil foi lançado em formato digital, e-book, disponibilizado de forma inteiramente gratuita. Este manual encontra-se disponível para download na página do Departamento de Arquivo Geral da UFSM.

Em uma abordagem de um ponto de vista técnico, o ICA-AtoM possui uma base forte, pois já atende aos padrões da descrição arquivística. Pode ser difundido facilmente, pois há muitas traduções oficiais ou já sendo oficializadas. Em relação aos custos ele supera expectativas, e é realmente uma ferramenta poderosa, com interface simples e profissional.

A sua continuidade no meio arquivístico se mostra pertinente e assim será confirmada com o tempo, pois o software está em processo de amadurecimento no que diz respeito ao seu desenvolvimento. É sem dúvida uma grande iniciativa do CIA, cabem agora as instituições, bem como compete aos arquivistas buscar conhecimento e implementar o ICA-AtoM institucionalmente.

REFERÊNCIAS

BRASIL. Conselho Nacional de Arquivos. NOBRADE: **Norma brasileira de descrição arquivística**. Rio de Janeiro: Arquivo Nacional, 2006.

CONSELHO INTERNACIONAL DE ARQUIVOS. ISAAR(CPF): **norma internacional de registro de autoridade arquivística para entidades coletivas, pessoas e famílias**. 2. ed. Rio de Janeiro: Arquivo Nacional, 2004.

_____. ISAD(G): **norma geral internacional de descrição arquivística**: segunda edição, Rio de Janeiro: Arquivo Nacional, 2000.



_____. ISDF: **norma internacional para descrição de funções**. Rio de Janeiro: Arquivo Nacional, 2008.

_____. ISDIAH: **norma internacional para descrição de instituições com acervo arquivístico**. Rio de Janeiro: Arquivo Nacional, 2009.

FONSECA, R. C; REZENDE, F. F. P. **Gestão arquivística: Normas e metodologias NOBRADE e ISAD(G)**. 2008.

HEDLUND, Dhion Carlos: **ICA-AtOM - Instalação, configuração e funcionalidades**. X SEMANA ACADÊMICA DO CURSO DE ARQUIVOLOGIA. Santa Maria: UFSM, 2012.

ICA-ATOM. **On-line User Manual: Glossary**. Disponível em: <<https://www.ica-atom.org/doc/Glossary>>. Acesso em 06 jul./2014.

PAVEZI, Neiva. **ICA-AtOM: manual do usuário em língua portuguesa**. Santa Maria, UFSM, 2013. Disponível em: <<http://w3.ufsm.br/dag/images/ICA-AtOM%20manual%20do%20usuário%20-%20PT%20BR.pdf>> Acesso em: 01 ago./2014.

SILVA, Edna Lúcia da; MENEZES, Estera Muszkat. **Metodologia da pesquisa e elaboração de dissertação**. – 4. Ed. rev. atual. – UFSC, Florianópolis, 2005.

UNIVERSIDADE FEDERAL DE SANTA MARIA (UFSM): patrimônio arquivístico. Online, disponível em: <www.ufsm.br/patrimonioarquivistico> Acesso em: 20 jul./2014.



PORTAL DE PERIÓDICOS DA UERJ NA PERSPECTIVA DO ACESSO LIVRE: UMA PROPOSTA

Ester Aparecida Lima de Souza

Técnico Administrativo Superior da Universidade do Estado do Rio de Janeiro - UERJ-
ester_limas@yahoo.com.br

RESUMO

Analisa o sítio de *Revistas Online* da Universidade do Estado do Rio de Janeiro (UERJ) onde são disponibilizados periódicos eletrônicos com o objetivo de propor um Portal em consonância com o movimento de acesso livre à informação. Traz uma reflexão a respeito do processo de comunicação científica estabelecido atualmente na Universidade. Lista os periódicos disponibilizados, suas formas de acesso e *softwares* utilizados. Descreve a importância de se implantar plataformas de acordo com *Open Archives Initiative* para a preservação da memória, ampliação do acesso, entre outras. Sugere que sejam verificados aspectos como acessibilidade, disponibilidade, interoperabilidade, padronização e preservação digital, adotando uma metodologia de levantamento bibliográfico sobre o assunto revistas eletrônicas, descrição dos periódicos eletrônicos disponibilizados pela Universidade e de suas plataformas tecnológicas com a finalidade de construir um ambiente digital que possa contemplar todos esses recursos de acesso aberto de forma centralizada. Espera-se contribuir efetivamente com a possibilidade de se ter um portal de periódicos na perspectiva do acesso livre com a capacidade de reunir, trocar informações científicas, facilitar as buscas de assuntos, aumentar a visibilidade das pesquisas e ampliar o acesso a esta produção.

Palavras-Chave: Periódicos eletrônicos. Acesso livre à informação. Portal de periódicos eletrônicos.

ABSTRACT

Analyzes the Journals Online Web site of the State University of Rio de Janeiro (UERJ) where electronic journals are available in order to propose a Portal in line with the movement of free access to information. Brings a reflection on the scholarly communication process currently established at the University. List electronic journals available, their forms of access and software used. Describes the importance of deploying platforms according to the Open Archives Initiative, to preserve the memory, expanding access, among others. Suggests that aspects are verified as accessibility, availability, interoperability, standardization and digital preservation, adopting a methodology of bibliographic electronic journals on the subject, description of electronic journals available from the University, and description of its technology platforms, with the aim of building a digital environment able to address all these open access resources centrally. Expected to contribute effectively to the possibility of having a journal portal from the perspective of free access with the



ability to gather, exchange scientific information, facilitate search subject, increase the visibility of research and expand access to this production.

Keywords: Electronic journals. Open Access. Scientific information

INTRODUÇÃO

Ao longo do desenvolvimento da humanidade pode-se eleger várias descobertas que funcionaram como marcos transformadores das relações sociais. Os principais foram sem dúvida: a imprensa, o rádio, o cinema, a televisão e a internet. A imprensa possibilitou ao homem maior divulgação do pensamento, perpetuou conceitos e estimulou revoluções. As ideias que até então eram difundidas através da oralidade, teve no livro, jornais, revistas e folhetos suportes para disseminação da informação. A internet surgiu basicamente de uma grande rede de computadores militares no período da Guerra Fria¹ e ninguém poderia imaginar que essa rede inicialmente criada com objetivos de controlar informações sigilosas, resultaria na web, um mundo conectado trocando informações com várias finalidades, inclusive a de entretenimento. A comunicação científica realizada através dos periódicos em formato papel, o modelo tradicional de divulgação de resultados de pesquisas, foi alterado e impactado quando passou a ter possibilidade de transmissão desses resultados online.

O movimento de acesso livre à informação científica foi resultante das transformações impulsionadas pelas Tecnologias de Informação e Comunicação (TICs), que vêm afetando o processo de disseminação da informação entre os pesquisadores. Iniciado na década de 90, preconiza o acesso de forma mais abrangente na web.

As universidades desfrutam de um papel estratégico nesse cenário, em constante transformação, representado pela inserção de novas tecnologias de informação. Mais do que apresentar soluções eletrônicas e digitais para resolução de problemas no âmbito da comunicação é necessário pensar muito sobre elas, avaliando todas as possibilidades e ofertas disponíveis para evitar duplicação de esforços e desperdício do dinheiro público.

A Universidade do Estado do Rio de Janeiro originou-se da Universidade do Distrito Federal com a união de quatro (4) Faculdades. São elas: Faculdade de Ciências Jurídicas, Faculdade de Ciências Econômicas e a de Filosofia, e Faculdade de Ciências e Letras. Quando a Capital mudou-se para Brasília passando a ser chamada de Universidade do Estado da Guanabara e em 1975 com a fusão dos estados da Guanabara e do Rio de Janeiro passou a ser conhecida como Universidade do Estado do Rio de Janeiro (UERJ). Atualmente oferece cursos de graduação nas áreas de Administração, Ciências Biológicas, Ciência da Computação, Ciências Contábeis, Ciências Econômicas, Ciências Sociais,

¹Guerra Fria segundo a Wikipédia é a designação atribuída ao período histórico de disputas estratégicas e conflitos indiretos entre os Estados Unidos e a União Soviética, compreendendo o período entre o final da Segunda Guerra Mundial (1945) e a extinção da União Soviética (1991), um conflito de ordem política, militar, tecnológica, econômica, social e ideológica entre as duas nações e suas zonas de influência. Fonte: Disponível em: <http://pt.wikipedia.org/wiki/Guerra_Fria>. Acesso em: 10 dez. 2013



Comunicação Social, Direito, Educação, Enfermagem, Engenharia, Filosofia, Física, Geografia, Geologia, História, Letras, Matemática, Medicina, Nutrição, Oceanografia, Odontologia, Psicologia, Serviço Social e Turismo, com habilitações em licenciaturas e bacharelados, e cursos de pós-graduação: especialização, mestrado e doutorado. São oferecidos em 30 unidades acadêmicas na cidade do Rio de Janeiro, nos campi do Maracanã, Duque de Caxias, Nova Friburgo, Resende, São Gonçalo e Teresópolis.

A Universidade do Estado do Rio de Janeiro (UERJ) tem sessenta anos de história e ao longo desse tempo vem prestando inúmeros serviços à população, desenvolvendo vários projetos de extensão, promovendo cultura e pesquisas, e sempre procurando investir em ações de divulgação de sua produção científica. Um exemplo dessas ações é a página eletrônica intitulada *Revistas Online*, <http://www.uerj.br/revistas/>, na qual a Instituição divulga 52 links de periódicos e segundo a própria descrição do sítio são adaptados à internet, disponíveis na web bastando um clique para que se tenha acesso a esses periódicos. Os periódicos são produzidos pelos Institutos que fazem parte da Universidade e a avaliação dos artigos é realizada por membros da comunidade científica de diversas instituições nacionais e internacionais.

O objetivo deste trabalho é propor uma reflexão sobre o processo de disponibilização de *Revistas Online* frente ao movimento de acesso livre à informação.

Os periódicos eletrônicos da Universidade do Estado do Rio de Janeiro sem dúvida têm facilitado o acesso à produção científica, porém cabe salientar que existem vários critérios como os de acessibilidade, disponibilidade, interoperabilidade, padronização dos assuntos presentes nesse universo de informações, autoarquivamento, modelos de comunicação científica, direitos autorais e tantos outros que poderiam ser analisados no intuito de promover discussões sobre a forma atual de divulgação dessas informações.

O periódico eletrônico, além de possibilitar a ampliação do acesso, é uma maneira de se adequar a dura realidade orçamentária das universidades públicas no que se refere à manutenção de assinaturas de revistas nos acervos em suas bibliotecas. A coexistência do formato de revistas em papel e em meio eletrônico, tem sido muito discutida, porém deve-se atentar para os problemas tecnológicos envolvidos, que precisam também de análises em suas políticas de divulgação:

A convivência das publicações impressas com as eletrônicas ainda ocorrerá por algum tempo, mas para que ambas sejam aceitas no meio científico os problemas de qualidade científica e os problemas tecnológicos - sejam de impressão, sejam de informática - precisam ser objeto de estudos, de políticas e de decisões, que incluam as revistas brasileiras como parte indispensável no processo de produção da ciência nacional. (STUMPF, 1998, p. 9)

Com a criação de *Revistas Online* a Universidade do Estado do Rio Janeiro, demonstra a preocupação de buscar soluções em se beneficiar das facilidades da internet contribuindo com a divulgação da informação, melhorando o acesso e a visibilidade desses artigos. Porém torna-se necessário a avaliação criteriosa desses *links*, pois o fato de estarem simplesmente disponibilizados na web, não significa que os mesmos estão acessíveis. O conceito de acessibilidade transcende o da disponibilidade, depende inclusive de todo conhecimento prévio que o usuário da informação deva ter para lidar com ela no meio tecnológico.

Muitos fatores fundamentam a utilização da comunicação eletrônica nas universidades, dentre eles: a pressão das agências de fomento, da própria instituição, o determinismo tecnológico e a necessidade de dar destaque às pesquisas desenvolvidas, aumentando dessa forma o acesso à informação, de acordo com Lawrence:

A disponibilidade online de um artigo pode não aumentar significativamente acesso e impacto, caso não haja serviços de busca eficientes e abrangentes, visto que um percentual substancial da literatura precisa ser indexado por esses serviços antes que os cientistas o considerem útil. (2001 apud COSTA, 2005 p.170)

Várias perguntas podem ser feitas como: todas essas revistas e seus artigos podem ser acessados e baixados em arquivos pdf ou foram digitalizadas no todo? Estão sendo produzidas exclusivamente em formato eletrônico, ou formato eletrônico e papel? Para se ter acesso ao seu conteúdo, é necessário em alguns casos a realização de cadastro? Outra questão a ser verificada é a da preservação digital, que deve ser vista com muito cuidado por uma instituição, principalmente de pesquisa e ensino que necessita preservar a memória.

Quanto à preservação, porém, o assunto é mais preocupante [...] a natureza descentralizada do provimento das revistas por redes não assegura que sua disponibilidade seja duradoura, e perguntam: onde estarão os textos das revistas eletrônicas daqui a 5, 10, 50 ou 100 anos? A preservação será uma tarefa apenas das bibliotecas nacionais? Quem organizará o arquivo das revistas eletrônicas? (STUMPF, 1996, p. 4)

Estudar criteriosamente o sítio *Revistas Online*, como estão sendo disponibilizadas, divulgadas e acessadas por sua comunidade acadêmica, é um campo rico em pesquisas e debates a respeito do movimento de acesso livre às publicações científicas brasileiras. Poderão surgir dessas análises questões como: políticas institucionais de informação, preservação da memória, acesso livre, interoperabilidade, direitos autorais, atuação de novos atores nesse processo como as editoras e agências de fomento à pesquisa.

Em países como Inglaterra, Alemanha e Estados Unidos, embora tenha predominância de editoras de periódicos científicos, os debates a respeito das políticas e diretrizes no que diz respeito ao acesso livre ao conhecimento científico estão mais avançados do que no Brasil em que vem sendo amplamente utilizado, por isso devemos ficar muito atentos, principalmente nas posições que estão sendo assumidas nesse tema por universidades, sociedades científicas e órgãos que financiam pesquisas. (BAPTISTA; COSTA; KURAMOTO; RODRIGUES, 2007).

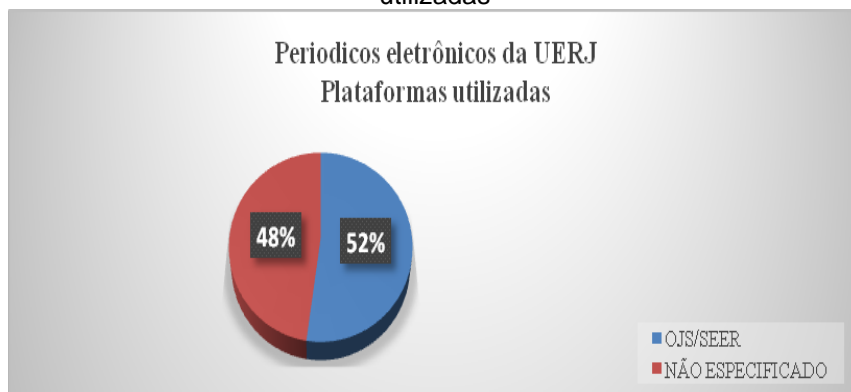
A Universidade do Estado do Rio de Janeiro (UERJ) tem a possibilidade de potencializar o desenvolvimento de suas pesquisas. Sendo assim é pertinente propor a criação de uma plataforma na qual seja possível centralizar, organizar, trocar informações e preservar sua coleção digital. Para o desenvolvimento de um portal com a perspectiva do acesso livre, será necessário, primeiramente, um estudo da real situação desses periódicos levando em conta alguns aspectos principais relacionados às categorias propostas por esse movimento como: filosofia aberta; acesso aberto (interoperabilidade); autoarquivamento; gestão de editoração, áreas de cobertura; revisores (comitês editoriais). Realizou-se um levantamento inicial com todos os títulos de periódicos disponibilizados em *Revistas Online*, a política de acesso e a plataforma utilizada por cada periódico como é demonstrado nos gráficos abaixo:

Gráfico 1- Revistas online da Universidade do Estado do Rio de Janeiro (UERJ) tipos de acesso.



Fonte: <http://www.uerj.br/revistas/> (gráfico elaborado pela autora)

Gráfico 2 – Revistas Online da Universidade do Estado do Rio de Janeiro (UERJ) plataformas utilizadas



Fonte: <http://www.uerj.br/revistas/> (gráfico elaborado pela autora)

Dos 52 *links* listados na página da UERJ, apenas 44 foram observados, pois 7 estavam com *links* inválidos. Dos 44 analisados, contatou-se que 42 periódicos (96 %) têm como política o acesso aberto e somente 2 o acesso fechado. A plataforma utilizada por 23 periódicos é o OJS (*Open Journal System*), um *software* desenvolvido pela Universidade British Columbia que foi customizado e introduzido no Brasil pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) como Sistema Eletrônico de Editoração de Revistas (SEER) para gestão de publicações eletrônicas, projetado para facilitar o acesso aberto, com facilidades de apresentação de artigos *online*. Sua principal vantagem é ser interoperável com o Dspace, um programa que vem sendo amplamente utilizado por universidades e instituições de pesquisa na criação de repositórios temáticos e institucionais. Os outros 21 periódicos não tinham a especificação da plataforma utilizada, necessitando de uma pesquisa mais profunda para obter a informação.

Esse levantamento inicial mostra que os periódicos na Universidade do Estado do Rio de Janeiro (UERJ) em sua maioria utiliza o SEER, um *software* que coopera para a criação de um portal de periódicos na perspectiva do acesso livre.

Um portal de acesso único, centralizado para os periódicos eletrônicos pode se tornar um meio mais confiável no momento das buscas de informações, pois o usuário terá uma visão global de todos os artigos produzidos. Atualmente esta tarefa é feita de forma árdua, tendo que percorrer *link* por *link* ao tentar fazer a busca em cada endereço eletrônico em separado quando este poderá disponibilizar a ferramenta de busca. As buscas por autor, assunto, título e palavras-chave ampliam as possibilidades do usuário encontrar a informação demandada.

A Universidade do Estado do Rio de Janeiro (UERJ) tem uma Rede de Bibliotecas, a Rede Sírius e a sua missão principal é a promoção do acesso à informação seguida do oferecimento de suporte as atividades de ensino, pesquisa e extensão, contribuindo para o desenvolvimento cultural, econômico e social do estado do Rio de Janeiro. Ela deve gerar debates, e consequentes avaliações do seu processo de comunicação científica, contribuindo com propostas de inserção no movimento de acesso livre. Sendo assim, este trabalho busca desenvolver um estudo minucioso da disponibilização e acesso dos periódicos eletrônicos na Universidade. Espera-se obter informações que sirvam de subsídios para a criação de um portal que permita o acesso de forma integrada, a toda literatura atualmente disponibilizada em Revistas *Online*.

OBJETIVOS: GERAL E ESPECÍFICO

Propor a criação de um portal de periódicos científicos na Universidade do Estado do Rio de Janeiro (UERJ) na perspectiva do acesso livre e:

- a) Realizar levantamento bibliográfico relacionado a portais de periódicos eletrônicos, e repositórios institucionais, especificamente quanto a integração de periódicos;
- b) Listar os periódicos eletrônicos disponibilizados pela Universidade, quanto a periodicidade, idade, políticas de acesso e plataforma tecnológica utilizada por cada periódico;

- c) Descrever e analisar os recursos de cada plataforma tecnológica;
- d) Propor um modelo que contemple a disponibilização destes recursos através de um único acesso.

REVISÃO DE LITERATURA

A comunicação científica tem sido objeto de estudo da Ciência da Informação há pelo menos três décadas, há predominância do enfoque quantitativo para explicar questões relacionadas com sua produção e o uso da literatura representado pelo estudo da citação. O trabalho de pesquisas é resultado de esforços mútuos, toda hipótese lançada, e até mesmo os resultados e suas comprovações, precisam ser avaliados por seus pares, é uma construção coletiva com olhares corporativos e cooperativos:

Os pesquisadores jamais percorrem sozinhos todos os degraus da cadeia lógico-indutiva, ao contrário, os percorrem em grupos e, enquanto dividem os frutos de seus esforços, estão também constantes, e invejosamente verificando, cada um, a contribuição do outro. E aqueles, para quem as publicações científicas são destinadas, não formam uma platéia passiva. (MUELLER, 2000, p.14)

Segundo Leite (2009) os periódicos científicos eletrônicos são considerados, um dos maiores avanços no processo de comunicação científica, porém a sua estruturação lógica, sofreu poucas alterações desde o seu aparecimento como modelo que se conhece hoje. O que realmente ocorreu foi a modernização no seu formato de publicação e o maior benefício nessa inovação foi sem dúvida o aumento das possibilidades de acesso aos seus conteúdos.

Foi na Grécia Antiga, berço da Civilização Ocidental que nasceu a troca de conhecimentos entre seus filósofos, os primeiros colégios invisíveis. Mas somente no século XVII que a Ciência consolidou seus saberes nos primeiros periódicos, científicos que até os dias atuais tem nesse formato sua aceitação, aprovação e difusão entre seus pares.

Segundo Kuramoto (2006) o americano Eugene Garfield criou o SCI (Science Citation Index) uma base de referência mundial com citações bibliográficas presentes em artigos de revistas, definidas segundo a classificação do seu impacto. Em consequência da alta valorização dos artigos indexados pela SCI e o reconhecimento de suas publicações, os editores promoveram a alta nos valores do preço de assinaturas de suas revistas, causando problemas na manutenção dos títulos de periódicos pelas bibliotecas e instituições de pesquisas, iniciando assim a crise dos periódicos.

É nesse contexto que surge o movimento de acesso livre à literatura científica mundial, uma resposta dos pesquisadores com várias manifestações, iniciadas na década de 90 e chegando ao seu apogeu na Convenção de Santa Fé, evento importante por ter lançado como principal proposta o modelo *Open Archives Initiative*, e também por ter discutido e consolidado uma linha de ação voltada para a definição de aspectos técnicos e organizacionais com a finalidade de possibilitar uma estrutura de publicação científica aberta, garantindo a interoperabilidade entre as informações disponibilizadas em meio eletrônico. Estabeleceram-se os seguintes requisitos para arquivos de *e-prints*: mecanismos de submissão; sistema de armazenamento a longo prazo; uma política de gestão para a preservação e submissão dos documentos; interface aberta que permita terceiros coletar os metadados dos respectivos arquivos; provedores de dados e de serviços.

As pesquisas científicas, aquelas que são realizadas nas universidades e financiadas com recursos públicos deveriam a princípio ser de acesso livre, não sendo isso o que ocorre realmente no sistema de comunicação científica tradicional, tendo o pesquisador ou qualquer outra pessoa interessada que pagar por aquilo que foi produzido com apoio do Estado, uma situação paradoxal segundo Kuramoto (2006).

MATERIAIS E MÉTODOS

Nessa seção serão descritas as etapas anteriormente citadas nos objetivos específicos, e os instrumentos necessários para realização das atividades propostas em cada fase do processo para se chegar a principal meta que é a criação de um Portal de Periódicos Eletrônicos da Universidade do Estado do Rio de Janeiro (UERJ).

Primeira etapa: Realização de um levantamento bibliográfico sobre portais e repositórios institucionais que tiveram a experiência de integrar a sua coleção digital, periódicos eletrônicos. Nessa etapa deverão ser utilizadas as redes de computadores, base de dados, catálogos e fontes de informação, com a finalidade de se obter uma relação de Repositórios Institucionais e Portais que possam servir de base para a criação de um modelo centralizado de disponibilização dos periódicos.

Segunda etapa: Descrever os periódicos eletrônicos da UERJ, título a título, elaborando uma tabela com os seguintes itens a serem analisados segundo critérios como:

- a) Periodicidade – anual, semestral, quinzenal, bimestral, quadrimestral, enfim em que período de tempo que os periódicos são divulgados na instituição.
- b) Políticas de acesso – esses periódicos estão com seus artigos disponibilizados em arquivos abertos e podem ser facilmente recuperados e baixados por qualquer computador; ou fechado, os artigos não estão disponíveis.
- c) Idade do periódico – em que ano cada título começou a ser disponibilizado.
- d) Plataforma tecnológica utilizada por cada periódico internet no processo de editoração eletrônica dessas Revistas.

Terceira etapa – Essa etapa utiliza critérios sugeridos segundo Kuramoto (2005) e pretende analisar e descrever os recursos de cada plataforma tecnológica, baseando-se nos requisitos desejáveis para um *software* de biblioteca digital:

- a) Portabilidade – É a capacidade de um pacote de *software* ser executado em diferentes ambientes operacionais;
- b) Flexibilidade quanto a definição do padrão de metadados – Possibilidade de oferecer ao usuário outros padrões de metadados;
- c) Uso de padrões de interoperabilidade – Facilidade de integração a outras iniciativas análogas, criação de redes. Uso de linguagem de marcação XML para facilitar a interoperabilidade entre os sistemas de bibliotecas digitais;
- d) Capacidade de tratamento de múltiplos formatos de documentos – Possibilidade de armazenar vários tipos de formatos como texto, imagem e som;
- e) Possuir interfaces ergonômicas e adaptativas – Para facilitar o seu uso por usuários menos especializados;
- f) Facilidade para estruturação dos documentos – poder considerar dois tipos de estruturação: no todo ou em partes (revista e artigo);
- g) Configuração dos procedimentos de indexação – Capacidade de definição dos metadados em campos de busca ou pontos de acesso, aos registros de uma biblioteca digital;
- h) Possuir módulo de formatação, de relatórios ou saídas – Capacidade de gerar relatórios ou configurar-se a apresentação de documentos.

Tipologia de pacotes de *software* para administração de bibliotecas digitais:

- a) Caráter genérico – Com capacidade de tratar, organizar, registrar e disseminar qualquer tipo de documento, por utilizar um padrão de metadados que tem compatibilidade com o padrão *Dublin Core* e podem ser utilizados na criação de qualquer tipo de bibliotecas digitais como por exemplo: *CDSware, E-prints, Nou-Rau, Phronesis, Fedora e Dspace*.
- b) Caráter específico – Atendem somente a determinados tipos de aplicações como exemplo o OJS, *Open Journal Systems*, construído pela Universidade Briths Columbia para administrar periódicos científicos eletrônicos e adaptado pelo IBICT como SEER, Sistema Eletrônico de Editoração de Revistas.

Quarta etapa – Deverá ser elaborado e proposto um modelo que consiga reunir em uma única plataforma os periódicos eletrônicos disponibilizados pela UERJ.



RESULTADOS PARCIAIS/FINAIS

A criação de um portal de acesso único dos periódicos eletrônicos na perspectiva do acesso livre poderá ser um meio de reúso das informações disponibilizadas tornando-se uma referência para a comunidade acadêmica e científica, principalmente se os arquivos forem interoperáveis com outros sistemas informacionais, maximizar o prestígio da instituição, valorizar a missão da universidade que se propõe a divulgar à população os resultados de suas pesquisas, tornando esse processo transparente. Pode vir a ser também, um instrumento eficaz para reduzir os custos de gestão da informação científica e uma ferramenta potente para auxiliar nas tomadas de decisões e também pode contribuir para melhorar na gestão dos direitos intelectuais e de propriedade da Instituição.

Um dos beneficiados com essa proposta será a própria comunidade científica que terá o fortalecimento do trabalho cooperativo em equipe e um reconhecimento tanto nacional, quanto internacional e o maior beneficiado será a população em geral tendo um amplo e irrestrito acesso à informação científica e tecnológica.

Considerações Parciais/Finais

A Constituição Federal brasileira, em seu artigo V, inciso XIV, assegura a todos o acesso à informação, portanto a sua democratização não pode ficar no plano das ideias, é um dever a ser cumprido e essa tarefa no âmbito do processo de comunicação científica, precisa ser avaliada e amplamente discutida por toda a sociedade. Com as novas tecnologias de informação, o que não faltam atualmente

são instrumentos e meios para disponibilizá-las, mas efetivamente isso não tem demonstrado ser suficiente. As universidades estão no centro desse debate cada vez mais necessário sobre o trabalho cooperativo em rede e da difusão e geração do conhecimento.

Devemos ter o cuidado de não reproduzir nas ofertas de serviços online, o mesmo caos informacional que a própria web retrata na fragmentação da informação.

Espera-se que esse projeto possa contribuir efetivamente com a ampliação da visibilidade dos periódicos eletrônicos disponibilizados pela Universidade do Estado do Rio de Janeiro (UERJ), pois de acordo com o levantamento realizado são 52 títulos, todos oriundos da comunicação em artigos resultantes de pesquisas científicas e devido a sua descentralização, dificulta o tratamento e organização dessas informações de forma padronizada, tornando difícil uma análise dos dados quantitativos de produção por parte dessas áreas de pesquisa e até mesmo saber quem produz o que na instituição. A centralização dessas coleções em um único acesso possibilitará um mapeamento e monitoramento dessas informações, obtendo indicadores mais precisos dessa produção científica. A facilidade de se ter uma plataforma que consiga organizar esses periódicos por áreas do conhecimento, pode servir como uma fonte segura de informação para as agências de fomento interessadas em investir em desenvolvimento de pesquisas em ciência e tecnologia.



A disponibilização das descobertas dos pesquisadores em ferramenta de acesso livre pode significar maior entrosamento e diálogo mais rápido entre seus pares.

O movimento de acesso livre à informação científica propõe estratégias claras e objetivas de ações como a implantação de repositórios institucionais e temáticos nas instituições de pesquisas e universidades. A criação deste portal pode ser o início para o desenvolvimento de um repositório institucional capaz de reunir a produção da UERJ, incluindo posteriormente outras tipologias como teses, dissertações e trabalhos publicados em eventos.

A preservação da memória deve ser vista como um dos fatores de suma importância dentro de uma instituição de pesquisa e ensino, e a criação de um portal de periódicos eletrônicos na perspectiva do acesso livre deve se preocupar em assegurar que os materiais digitais continuem acessíveis e disponíveis ao longo do tempo.

REFERÊNCIAS

ARAYA, Elizabeth Roxana Mass; VIDOTTI, Silvana Aparecida BorsettiGregorio. Direito autoral e tecnologia de informação e comunicação no contexto da produção, uso e disseminação da informação: um olhar para as licenças *creative commons*.

Informação & Sociedade: estudos, João Pessoa, v. 19, n.3, p.39-51, set./dez. 2009. Disponível em:

<http://eprints.rclis.org/14443/1/10_ARAYA_VIDOTTI_DIREITO_AUTORAL%5B1%5D.pdf>. Acesso em: 9 ago. 2013.

BAPTISTA, Ana Alice et al. Comunicação científica: o papel da open archives initiative no contexto do acesso livre. **Encontros Bibli:** revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, n. esp., p.1-17, 1º sem. 2007. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2007v12nesp1p1/435>>. Acesso em: 8 ago. 2013.

BRASIL. Constituição (1988). Constituição [da] Republica Federativa do Brasil. Brasília, DF: Senado Federal.

BRODY, Tim; HARNAD, Steva. **The research impact cycle**. PowerPoint presentation. Disponível em: <www.ecs.soton.ac.uk/~harnad/Temp/self-archiving.ppt>. Acesso em: 20 nov. 2013.

COSTA, Sely Maria de Souza. Mudanças no processo de comunicação científica: o impacto do uso de novas tecnologias. In: MUELLER, Suzana Pinheiro Machado; PASSOS, Edilenice. (Org.). **Comunicação científica**. Brasília: Departamento de Ciência da Informação da Universidade de Brasília, 2000. p. 95-105. Disponível em:

<http://repositorio.unb.br/bitstream/10482/1443/1/CAPITULO_MudancaProcessoComunicacao.pdf>. Acesso em: 15 jul. 2013.



COSTA, Sely Maria de Souza. O novo papel das tecnologias digitais na comunicação científica. In: MARCONDES, C.H.(Org.). **Bibliotecas digitais: saberes e práticas**. Salvador: EDUFBA; Brasília: IBICT, 2005 p. 165-183.

GARVEY, William D., GRIFFITH, Belver G. Communication and information processing within scientific disciplines: empirical findings. In: GARVEY, William D. **Communication: the essence of science**. Oxford :Pergamon Press, p. 127-147, 1979.

KURAMOTO, Helio. Ferramentas de *software* livre para bibliotecas digitais. In: MARCONDES, C.H.(Org.). **Bibliotecas digitais: saberes e práticas**. Salvador: EDUFBA; Brasília: IBICT, 2005 p. 145-162.

KURAMOTO, Hélio. Informação científica: proposta de um novo modelo para o Brasil. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 91-102, 2006.

LAWRENCE, Steve. **Free online availability substantially increases a paper`s impact**. Naturewebdebates. Disponível em:
<<http://www.nature.com/nature/debates/debates/e-access/Articles/lawrence.html>>
Acesso em: 03 jun. 2013.

LEITE, Fernando Cesar Lima. **Como ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto**. Brasília: IBICT, 2009. p. Disponível em: <http://eprints.rclis.org/13776/1/RI_-_Fernando_Leite.pdf>. Acesso em: 7 out. 2013.

LEITE, Fernando Cesar Lima. Informação científica: proposta de um novo modelo para o Brasil. **Ciência da informação**, Brasília, v.35, n.2, p.91-102, maio./ago. 2006. Disponível em: < <http://www.scielo.br/pdf/ci/v35n2/a10v35n2.pdf>>. Acesso em: 15 ago 2013.

MORIGI, Valdir; SANTIN, Dirce Maria. Reflexões sobre os valores do movimento *software* livre. **Informação & Informação**, Londrina, v. 12, n.1, jan./jun., 2009. Disponível em:
<<http://www.uel.br/revistas/uel/index.php/informacao/article/viewArticle/1746>>
Acesso em: 10 jul. 2013.

MUELLER, Suzana P. M.; PASSOS, Edilenice J. L.. As questões da comunicação científica e a ciência da informação. In: MUELLER, Suzana P. M.; PASSOS, Edilenice J. L. (Org.). **Comunicação científica**. Brasília: Ciência da Informação, 2000. p. 13-22. Disponível em: < <http://repositorio.unb.br/handle/10482/1444>>. Acesso em: 14 jul. 2013.



MUELLER, Suzana Pinheiro Machado. A comunicação científica e o movimento de acesso livre ao conhecimento, **Ciência da informação**, Brasília, v. 35, n.2, p.27-38, maio/ago., 2006. Disponível em: <<http://www.scielo.br/pdf/ci/v35n2/a04v35n2.pdf>>. Acesso em: 4 set. 2013.

STUMPF, Ida Regina Chitto. O passado e futuro das revistas científicas. **Ciência da informação**,: v. 25, n. 3, 1996. Disponível em:<<http://revista.ibict.br/ciinf/index.php/ciinf/article/viewPDFInterstitial/463/422%26gt>>. Acesso em: 6 de jul. 2013.

STUMPF, Ida Regina Chitto. Reflexões sobre as revistas brasileiras. **Intexto**, Porto Alegre: UFRGS, v. 1, n. 3, p. 1-10, jan./jun, 1998. Disponível em: <<http://seer.ufrgs.br/index.php/intexto/article/viewFile/3369/3953>>. Acesso em: 26 set. 2013.



Visibilidade

Métricas de impacto e visibilidade dos documentos em acesso aberto, Divulgação, sensibilização dos usuários, capacitação



OS INVESTIGADORES E A SUA RELAÇÃO COM O ACESSO ABERTO À PRODUÇÃO CIENTÍFICA: O CASO DO ISCTE-IUL

Maria João Amante¹; Teresa Segurado¹

¹ISCTE-IUL – Instituto Universitário de Lisboa

INTRODUÇÃO

A presente comunicação resulta de uma atividade desenvolvida pelos Serviços de Informação e Documentação e pelo Gabinete de Apoio à Investigação do ISCTE-IUL no domínio do Acesso Livre ao Conhecimento e surge na sequência de um estudo semelhante realizado, em 2012, pelos Serviços de Documentação da Universidade do Minho, junto dos investigadores que, em Portugal, participam em projetos financiados pela Fundação para a Ciência e a Tecnologia (FCT) e pelo 7º Programa Quadro da Comissão Europeia no contexto das atividades desenvolvidas nos projetos OpenAIRE, MedOAnet e OpenAIREplus.

Constituiu objetivo deste trabalho recolher informações sobre a perceção, opinião e prática dos investigadores do ISCTE-IUL relativamente a cinco dimensões no que respeita ao acesso aberto aos resultados da investigação científica:

- a) o nível de conhecimento, apreciação e atitude sobre o princípio do acesso aberto aos resultados da investigação;
- b) o nível de conhecimento sobre as políticas e as diretrizes da Comissão Europeia em matéria de acesso aberto;
- c) o nível de cumprimento relativamente à política do ISCTE-IUL em matéria de acesso aberto;
- d) a opinião e a atitude perante a eventual definição de um mandato por parte das organizações públicas portuguesas financiadoras de ciência requerendo a disponibilização em acesso aberto das publicações resultantes de financiamento público;
- e) a prática de publicação científica e de disponibilização dos resultados em acesso aberto através de repositórios e/ou revistas.

Tomando em consideração o âmbito da Conferência centramo-nos na apresentação dos resultados obtidos nas três últimas dimensões por os considerarmos pertinentes para a análise das questões relativas à visibilidade da produção científica através do recurso ao Acesso Aberto.



METODOLOGIA

A recolha de informação foi feita através da aplicação de um inquérito por questionário estruturado em quatro grupos distintos composto por um total de 37 questões. O primeiro grupo caracteriza o perfil do investigador/docente participante no estudo. O segundo grupo caracteriza a participação em projetos com financiamento pela FCT – Fundação para a Ciência e a Tecnologia e pelo 7º Programa-Quadro da

Comissão Europeia. O terceiro grupo de questões identifica o nível de conhecimento, opinião e atitude sobre o princípio do acesso aberto aos resultados da investigação. O quarto grupo caracteriza as práticas de publicação e de acesso aberto dos investigadores/docentes. Conforme referido anteriormente, nesta comunicação analisamos os dados obtidos nos grupos de questões um, três e quatro.

Foi enviado um *e-mail* a convidar todos os docentes e investigadores do ISCTE-IUL para a participação neste estudo tendo sido, igualmente, realizados contactos personalizados no sentido de obter um maior número de respostas.

O questionário foi aplicado via *Web*, através do *software* LimeSurvey e cujo período para resposta decorreu de 1 a 19 de julho 2013.

O tratamento estatístico dos dados recolhidos foi realizado com o auxílio do IBM SPSS Statistics, versão 20, e do Microsoft Excel para a construção dos gráficos.

Para o tratamento dos dados utilizou-se quer a estatística descritiva quer a indutiva admitindo-se, na última, uma probabilidade de erro (nível de significância) até 10% através da utilização de testes de hipóteses não paramétricos (*Teste de independência do Qui-quadrado, Teste de Mann-Withney e Teste de Kruskal-Wallis*).

As análises realizadas foram de âmbito univariado e, sobretudo, bivariado, para tentarmos perceber se as respostas apuradas variavam segundo as características dos respondentes bem como de acordo com a sua participação ou não em projetos.

Resultados da análise dos dados do inquérito por questionário

Responderam ao inquérito por questionário 246 pessoas porém apenas 143 respostas puderam ser consideradas para efeitos de análise dos dados apurados por estarem completas.

Caracterização

No que respeita à *Carreira Profissional* houve uma maior adesão por parte dos *Investigadores* com 65,0% (n=93) de respostas face a 35,0% (n=50) dos *Docentes*.

Quando analisada a participação por *Departamento* constatamos que o Departamento que maior percentagem apresentou foi o de *Ciências e Tecnologias da Informação* com 16,7% seguido do de *Sociologia* com 14,6%. Apenas os *Departamentos de Contabilidade* e de *História* não contaram com a participação de nenhum membro.



No que toca à participação das *Unidades de Investigação* verificamos que foi o *CIES* que mais participou (31,5%) seguido do *Cis-IUL* com 26,1%. A única Unidade de Investigação que não registou qualquer participação foi o *IT-IUL*.

Importa referir que, quer no que respeita aos Departamentos quer às Unidades de Investigação, os que mais participaram neste estudo foram os que, igualmente, mais colaboram no desenvolvimento do Repositório ISCTE-IUL.

Ainda no que respeita a caracterização dos respondentes e relativamente ao *Sexo* foram as *Mulheres* que mais participaram neste estudo com 63,6% das respostas.

No que toca à *Idade* verificamos que a faixa etária com maior representação situa-se nos *31 aos 40 anos* com 39,2% de respostas, grupo este que abrange perto de metade dos investigadores que respondeu ao questionário.

No ISCTE-IUL, os docentes e investigadores estão organizados por Escolas, Departamentos e Unidades de Investigação de acordo com a seguinte estrutura:

Tabela 1 Estrutura organizacional do ISCTE-IUL

ISCTE BUSINESS SCHOOL - IBS	Escola de Sociologia e Políticas Públicas - ESPP	Escola de Ciências Sociais e Humanas - ECSH	Escola de Tecnologias e Arquitetura - ISTA
Departamento de Contabilidade	Departamento de Métodos de Pesquisa Social	Departamento de Economia Política	Departamento de Matemática
Departamento de Economia	Departamento de Ciência Política e Políticas Públicas	Departamento de Antropologia	Departamento de Arquitetura e Urbanismo
Departamento de Finanças	Departamento de Sociologia	Departamento de Psicologia Social e das Organizações	Departamento de Ciências e Tecnologias da Informação
Departamento de Marketing, Operações e Gestão Geral	Departamento de História	CIS-IUL – Centro de Investigação e Intervenção Social	ADETTI
Departamento de Métodos Quantitativos para Gestão e Economia	CEI-IUL – Centro de Estudos Internacionais	Dinâmia/CET – Centro de Estudos sobre a mudança socioeconómica e o território	IT-IUL - Delegação do Instituto de Telecomunicações
Departamento de Recursos Humanos e Comportamento Organizacional	CEHC-IUL – Centro de Estudos de História Contemporânea	CRIA-IUL – Pólo do Centro em rede de Investigação em antropologia	
Business Research Unit - BRU-IUL	CIES-IUL – Centro de Investigação e Estudos em Sociologia		

Sobre o Acesso Livre

O terceiro grupo, sobre o Acesso Livre, era constituído por 16 questões e procurou aferir os conhecimentos, as opiniões e as perceções dos investigadores e docentes do ISCTE-IUL sobre o acesso aberto, barreiras no acesso à informação científica, mandatos e políticas institucionais e de organismos financiadores.

Quando questionados se concordam com a afirmação “*Não existe qualquer problema com o acesso à informação científica em Portugal*” 87,6% dos respondentes afirmaram *Discordar* e *Discordar plenamente* da mesma. Curiosamente a percentagem de respondentes que escolheu a opção *Concordo* é exatamente igual à dos que escolheram a opção *Discordo plenamente* (12,6%), o que nos leva a crer que existe um conjunto de investigadores que consegue aceder mais facilmente à informação científica enquanto outros sentem uma grau de dificuldade muito mais elevado, o que pode, de certo modo, ser muito influenciado pela área científica de cada indivíduo pois a forma como as publicações científicas são disponibilizadas bem como, por exemplo, o tempo de embargo das revistas científicas varia de área para área.

Quando repartidos segundo a *Carreira Profissional* os valores acima representados mostram-nos que quer os *Docentes* quer os *Investigadores* são de opinião que existem de facto problemas no acesso à informação científica em Portugal sendo essa opinião mais acentuada no grupo dos *Investigadores* pois 79,6% *Discordam* e *Discordam plenamente* face a 74,0% dos *Docentes*.

No que respeita ao Departamento são os *Docentes* pertencentes a *Departamentos integrados na ECSH* que mais discordam desta afirmação (85,7% *Discordam* e *Discordam plenamente*). Já no que respeita à *Unidade de Investigação* são os membros daquelas que pertencem à *ETA / EG* cujos *Investigadores* mais discordam. No que refere ao *Sexo* são claramente as *Mulheres* que mais discordam (83,5% *Discordam* e *Discordam plenamente*). Relativamente à *Idade* é o escalão que se encontra entre os 31 e os 40 anos que mais discorda da afirmação apresentada (87,5% *Discordam* e *Discordam plenamente*) sendo esta a única relação estatisticamente significativa, ou seja, apenas a *Idade* influencia o grau de concordância. De referir que esta relação é inversa, ou seja, quanto mais novos mais discordam.

Tabela 2 - Nível de concordância com a afirmação: “*Não existe qualquer problema com o acesso à informação científica em Portugal*” de acordo com as variáveis de caracterização (% em linha)

		Concordo plenamente		Concordo		Sem opinião		Discordo		Discordo plenamente	
		n	%	n	%	n	%	n	%	n	%
Carreira Profissional	Docente	1	2,0	11	22,0	1	2,0	32	64,0	5	10,0
	Investigador	2	2,2	7	7,5	10	10,8	61	65,6	13	14,0
Departº/Escola	ECSH	1	7,1	1	7,1	0	,0	9	64,3	3	21,4
	ETA / EG	0	,0	5	29,4	1	5,9	10	58,8	1	5,9
	ESPP	0	,0	5	29,4	0	,0	11	64,7	1	5,9
Unid. de Invest. / Escola	ECSH	2	4,4	1	2,2	6	13,3	29	64,4	7	15,6
	ETA / EG	0	,0	0	,0	0	,0	4	80,0	1	20,0
	ESPP	0	,0	6	14,3	4	9,5	27	64,3	5	11,9
Sexo	Feminino	3	3,3	6	6,6	6	6,6	65	71,4	11	12,1
	Masculino	0	,0	12	23,1	5	9,6	28	53,8	7	13,5
Idade	≤ 30	1	3,4	4	13,8	1	3,4	15	51,7	8	27,6
	31-40	0	,0	3	5,4	4	7,1	43	76,8	6	10,7
	41-50	1	2,6	8	20,5	6	15,4	23	59,0	1	2,6
	≥ 51	1	5,3	3	15,8	0	,0	12	63,2	3	15,8

Carreira Profissional (M-WZ=-1,162; p=0,245) / Departamento/Escola (KW= 2,500; p=0,287) / Unidade de Investigação / Escola (KW= 1,507; p=0,471) / Sexo (M-WZ=-1,510; p=0,131) / Idade (KW= 10,530; p=0,015)

Quando inquiridos sobre as potenciais barreiras de acesso às publicações científicas a maioria dos participantes neste estudo considera os *Elevados preços dos artigos/assinaturas das revistas científicas* como a principal barreira para o acesso às publicações científicas pois 87,8% considera como *Importante e Muito importante* este fator. Em segundo lugar surgem as *Insuficientes políticas, estratégias e estruturas de acesso às publicações científicas* (83,0%).

Quando pedida a opinião sobre o princípio da disponibilização em acesso aberto das publicações científicas resultantes de projetos financiados por programas públicos a maioria dos participantes neste estudo é de opinião favorável à disponibilização em acesso aberto das publicações científicas resultantes de projetos financiados por programas públicos com cerca de 94,4% dos respondentes que afirmam *Concordar plenamente e Concordar* com esta regra.

Quando distribuídos estes valores pelas variáveis de caracterização é possível apurar que os que mais concordam com este princípio são os *Investigadores* (68,6%), os *Docentes de Departamentos integrados na ESPP* (42,6%) e os *membros das Unidades de Investigação pertencentes à ECSH e ESPP* ambas com 44,9%. Quanto ao *Sexo* são as *Mulheres* que mais concordância revelam com esta regra (60,3%) e na *Idade* são os participantes que se encontram entre os 31 e os 40 anos. Apesar das diferenças assinaladas apenas as encontradas na variável *Carreira Profissional* são estatisticamente significativas.

Tabela 3 Opinião em relação ao princípio da disponibilização em acesso aberto das publicações científicas resultantes de projetos financiados por programas públicos de acordo com as variáveis de caracterização

		Concordo plenamente		Concordo		Sem opinião		Discordo		Discordo plenamente	
		n	%	n	%	n	%	n	%	n	%
Carreira Profissional	Docente	35	70,0	12	24,0	1	2,0	2	4,0	0	,0
	Investigador	77	82,8	11	11,8	4	4,3	1	1,1	0	,0
Deptº / Esc	ECSH	12	85,7	1	7,1	0	,0	1	7,1	0	,0
	ETA / EG	10	58,8	5	29,4	1	5,9	1	5,9	0	,0
	ESPP	12	70,6	5	29,4	0	,0	0	,0	0	,0
Unid / Esc	ECSH	37	82,2	7	15,6	0	,0	1	2,2	0	,0
	ETA / EG	4	80,0	1	20,0	0	,0	0	,0	0	,0
	ESPP	35	83,3	3	7,1	4	9,5	0	,0	0	,0
Sexo	Feminino	72	79,1	16	17,6	1	1,1	2	2,2	0	,0
	Masculino	40	76,9	7	13,5	4	7,7	1	1,9	0	,0
Idade	≤ 30	25	86,2	4	13,8	0	,0	0	,0	0	,0
	31-40	45	80,4	8	14,3	2	3,6	1	1,8	0	,0
	41-50	26	66,7	9	23,1	3	7,7	1	2,6	0	,0
	≥ 51	16	84,2	2	10,5	0	,0	1	5,3	0	,0

Carreira Profissional (M-WZ=-1,695; p=0,090) / Departamento/Escola (KW= 2,495; p=0,287) / Unidade de Investigação / Escola (KW= 0,008; p=0,996) / Sexo (M-WZ=-0,456; p=0,648) / Idade (KW= 4,925; p=0,177)

Quando analisadas estas opiniões face à participação em projetos financiados pela FCT – Fundação para a Ciência e a Tecnologia, no âmbito do 7º Programa-Quadro da CE ou noutra tipo de projetos percebemos que apenas a participação em projetos financiados no âmbito do 7º Programa-Quadro da CE influencia a opinião dos participantes neste estudo ($M-WZ=-2,143$; $p=0,032$).

Tabela 4 Opinião em relação ao princípio da disponibilização em acesso aberto das publicações científicas resultantes de projetos financiados por programas públicos de acordo com a participação em projetos (% em linha)

		Concordo plenamente		Concordo		Sem opinião		Discordo		Discordo plenamente	
		n	%	n	%	n	%	n	%	n	%
Participa ou participou em projetos financiados pela FCT – Fundação para a Ciência e a Tecnologia?	Sim	79	76,0	20	19,2	2	1,9	3	2,9	0	,0
	Não	33	84,6	3	7,7	3	7,7	0	,0	0	,0
Participa ou participou em projetos financiados no âmbito do 7º Programa-Quadro da Comissão Europeia?	Sim	31	91,2	3	8,8	0	,0	0	,0	0	,0
	Não	81	74,3	20	18,3	5	4,6	3	2,8	0	,0
Participa ou participou noutros projetos?	Sim	62	75,6	16	19,5	2	2,4	2	2,4	0	,0
	Não	50	82,0	7	11,5	3	4,9	1	1,6	0	,0

Participação em projetos financiados pela FCT ($M-WZ=-0,994$; $p=0,320$) / Participação em projetos financiados no âmbito do 7º Programa-Quadro da CE ($M-WZ=-2,143$; $p=0,032$) / Participação noutros projetos ($M-WZ=-0,813$; $p=0,416$)

Quando questionados sobre a política europeia de acesso aberto¹, 58,3% dos respondentes a este questionário *Concorda* e *Concorda plenamente* com a política geral da Comissão Europeia relativamente ao acesso aberto, sendo que 40,5% manifestou não ter opinião.

Foi igualmente possível apurar que apenas a participação em projetos financiados no âmbito do 7º PQ ($M-WZ=-2,006$; $p=0,045$) influencia o grau de concordância dos participantes neste estudo face à política europeia de acesso aberto de acordo.

Quando questionados sobre a possibilidade de uma eventual política/mandato de acesso aberto da FCT – Fundação para a Ciência e a Tecnologia, requerendo o acesso aberto às publicações produzidas pelos projetos e bolsas que financia a maioria dos respondentes *Concorda* e *Concorda plenamente* (89,5%).

¹ A Comissão Europeia tem atualmente duas diretrizes políticas de acesso aberto que visam assegurar que os resultados da investigação financiada pelos cidadãos da EU sejam disponibilizados em acesso aberto para a população em geral. Em dezembro de 2007, o Conselho Europeu de Investigação publicou as diretrizes para o Open Access, e em agosto de 2008 a Comissão Europeia lançou o projeto-piloto Open Access do 7º PQ.

Analisando estes valores segundo as variáveis de caracterização é possível verificar que são os Investigadores que revelam um maior nível de concordância (91,4% *Concordam e Concordam plenamente*). No tocante ao *Departamento* e à *Unidade de Investigação* são quem pertence aos que integram a ECSH que mais concordam a atingir em ambos os caso valores próximos dos 100%. No que refere ao Sexo são as *Mulheres* que mais concordam com uma eventual política/mandato de acesso aberto da FCT (93,4% *Concorda e Concorda plenamente*). Na *Idade* são também os investigadores e docentes mais novos (≤ 30) que apresentam níveis de concordância mais elevados (96,5% *Concordam e Concordam plenamente*).

Das relações assinaladas apenas a pertença departamental revelou evidência estatística, ou seja, o departamento a que pertence influencia a opinião sobre uma eventual política/mandato de acesso aberto da FCT – Fundação para a Ciência e a Tecnologia que requeresse o acesso aberto às publicações produzidas pelos projetos e bolsas que financia.

Tabela 5 Opinião sobre uma política de acesso aberto da FCT segundo as variáveis de caracterização (% em linha)

		Concordo plenamente		Concordo		Sem opinião		Discordo		Discordo plenamente	
		n	%	n	%	n	%	n	%	n	%
Carreira	Docente	35	70,0	8	16,0	5	10,0	2	4,0	0	,0
	Profissional Investigador	59	63,4	26	28,0	3	3,2	4	4,3	1	1,1
Deptº / Esc	ECSH	14	100,0	0	,0	0	,0	0	,0	0	,0
	ETA / EG	9	52,9	2	11,8	4	23,5	2	11,8	0	,0
	ESPP	12	70,6	4	23,5	1	5,9	0	,0	0	,0
Unid / Esc	ECSH	25	55,6	18	40,0	0	,0	2	4,4	0	,0
	ETA / EG	4	80,0	0	,0	0	,0	1	20,0	0	,0
	ESPP	30	71,4	8	19,0	3	7,1	0	,0	1	2,4
Sexo	Feminino	63	69,2	22	24,2	2	2,2	3	3,3	1	1,1
	Masculino	31	59,6	12	23,1	6	11,5	3	5,8	0	,0
Idade	≤ 30	19	65,5	9	31,0	0	,0	1	3,4	0	,0
	31-40	35	62,5	13	23,2	3	5,4	4	7,1	1	1,8
	41-50	24	61,5	10	25,6	4	10,3	1	2,6	0	,0
	≥ 51	16	84,2	2	10,5	1	5,3	0	,0	0	,0

Carreira Profissional (M-WZ=-0,525; p=0,599) / Departamento/Escola (KW= 9,260; p=0,010) / Unidade de Investigação / Escola (KW= 1,760; p=0,415) / Sexo (M-WZ=-1,402; p=0,161) / Idade (KW= 3,597; p=0,308)

Quanto à influência da participação em qualquer tipo de projetos no grau de concordância relativamente a uma política de acesso aberto da FCT, e da análise dos dados apurados, não foi encontrada nenhuma evidência estatística de que a mesma exista, ou seja, a participação em projetos financiados pela FCT, no âmbito do 7º PQ da CE ou noutros não exerceu qualquer influência na opinião dos respondentes.

Tabela 6 Opinião sobre uma política de acesso aberto da FCT de acordo com a participação em projetos (% em linha)

		Concordo plenamente		Concordo		Sem opinião		Discordo		Discordo plenamente	
		n	%	n	%	n	%	n	%	n	%
Participa ou participou em projetos financiados pela FCT – Fundação para a Ciência e a Tecnologia?	Sim	70	67,3	25	24,0	4	3,8	5	4,8	0	,0
	Não	24	61,5	9	23,1	4	10,3	1	2,6	1	2,6
Participa ou participou em projetos financiados no âmbito do 7º Programa-Quadro da Comissão Europeia?	Sim	20	58,8	10	29,4	2	5,9	2	5,9	0	,0
	Não	74	67,9	24	22,0	6	5,5	4	3,7	1	,9
Participa ou participou noutros projetos?	Sim	57	69,5	15	18,3	5	6,1	4	4,9	1	1,2
	Não	37	60,7	19	31,1	3	4,9	2	3,3	0	,0

Participação em projetos financiados pela FCT (M-WZ=-0,800; p=0,423) / Participação em projetos financiados no âmbito do 7º PQ da CE (M-WZ=-0,911; p=0,362) / Participação noutros projetos (M-WZ=-0,782; p=0,434)

Já se verificarmos a correlação entre a opinião sobre uma eventual política de *acesso aberto da FCT* e a *opinião de quem participou sobre a política europeia de acesso aberto* e a *opinião sobre o princípio da disponibilização em acesso aberto das publicações científicas resultantes de projetos financiados por programas públicos* é possível confirmar a existência de uma correlação positiva, embora fraca, entre estas variáveis.

É, igualmente, possível apurar que o nível de concordância com a *política europeia de acesso aberto* e com o *princípio da disponibilização em acesso aberto das publicações científicas* influencia a opinião dos participantes neste estudo, sobre uma *eventual política/mandato de acesso aberto da FCT* (*Opinião sobre a política europeia de acesso aberto* (KW= 20,080; p=0,000 / *Opinião sobre o princípio da disponibilização em acesso aberto das publicações científicas* KW= 21,495; p=0,000).

Tabela 7 Correlação entre opinião sobre política/mandato de acesso aberto da FCT e política europeia de acesso aberto; princípio da disponibilização em acesso aberto das publicações científicas

		Opinião sobre a política europeia de acesso aberto		Opinião sobre o princípio da disponibilização em acesso aberto das publicações científicas	
Rho de spearman	Opinião sobre uma eventual política/mandato de acesso aberto da FCT	Coefficiente de correlação	,343**		,386**

Quando questionados sobre a *reação perante um mandato de acesso aberto à produção científica implementado pela FCT* a grande maioria *Cumpriria integralmente e concordaria* (85,3%). Estes resultados estão em linha com os obtidos relativamente à disponibilização em acesso aberto das publicações científicas resultantes de projetos financiados por programas públicos em que cerca de 94,4% dos respondentes afirma *Concordar plenamente e Concordar* com esta regra.

Quando questionados sobre o *Conhecimento relativamente à existência no ISCTE-IUL, de uma política de acesso aberto à produção científica* apenas 25,9% respondeu positivamente.

Distribuídos estes valores pelas variáveis de caracterização é possível apurar que apesar das diferenças verificadas apenas no que respeita à *Idade* ($KW= 10,770$; $p=0,013$) e à *Unidade de Investigação* ($KW= 10,797$; $p=0,005$) as mesmas são mais acentuadas e estatisticamente significativas, ou seja, são estas duas características que influenciam o conhecimento da existência de uma política de acesso aberto à produção científica no ISCTE-IUL.

Tabela 8 Conhecimento da existência no ISCTE-IUL, de uma política de acesso aberto à produção científica segundo as variáveis de caracterização

		n	%
Carreira Profissional	Docente	16	45,7
	Investigador	19	54,3
Deptº / Esc	ECSH	5	31,3
	ETA / EG	4	25,0
	ESPP	7	43,8
Unid / Esc	ECSH	4	21,1
	ETA / EG	0	,0
	ESPP	15	78,9
Sexo	Feminino	25	71,4
	Masculino	10	28,6
Idade	≤ 30	4	11,4
	31-40	14	40,0
	41-50	7	20,0
	≥ 51	10	28,6

Carreira Profissional ($X^2=2,355$; $p=0,125$) / Departamento/Escola ($KW= 1,216$; $p=0,545$) / Unidade de Investigação / Escola ($KW= 10,797$; $p=0,005$) / Sexo ($X^2=1,216$; $p=0,270$) / Idade ($KW= 10,770$; $p=0,013$)

Verificando a possível influência da participação em qualquer tipo de projetos de investigação no grau de conhecimento da existência no ISCTE-IUL, de uma política de acesso aberto à produção científica, e da análise dos dados apurados, foram encontradas evidências estatísticas de que a participação em projetos financiados pela FCT ($X^2=3,941$; $p=0,047$) e noutros projetos ($X^2=3,759$; $p=0,053$) influencia o nível de conhecimento da existência no ISCTE-IUL, de uma política de acesso aberto à produção científica.

Tabela 9 Conhecimento da existência no ISCTE-IUL, de uma política de acesso aberto à produção científica de acordo com a participação em projetos (% em linha)

		n	%
Participa ou participou em projetos financiados pela FCT – Fundação para a Ciência e a Tecnologia?	Sim	30	28,8%
	Não	5	12,8%
Participa ou participou em projetos financiados no âmbito do 7º Programa-Quadro da Comissão Europeia?	Sim	5	14,7%
	Não	30	27,5%
Participa ou participou noutros projetos?	Sim	25	30,5%
	Não	10	16,4%

Participação em projetos financiados pela FCT ($X^2=3,941$; $p=0,047$) / Participação em projetos financiados no âmbito do 7º PQ da CE ($X^2=2,303$; $p=0,129$) / Participação noutros projetos ($X^2=3,759$; $p=0,053$)

Relativamente ao *nível de cumprimento da política de acesso aberto à produção científica do ISCTE-IUL* do conjunto que afirma conhecer a política de acesso aberto à produção científica do ISCTE-IUL, 80,0% *Cumpr integralmente e concorda* com a mesma e 11,4% afirma que apenas *Cumpr parcialmente*. De salientar que uma pequena parcela dos respondentes afirma que *Não cumpre de todo* (5,7%).

Quando repartidos estes valores pelas diferentes variáveis de caracterização são os *Investigadores*, docentes dos *Departamentos pertencentes à ECSH*, investigadores das *Unidades de Investigação pertencentes à ECSH, Homens e Mulheres* em igual percentagem e os participantes com *idades iguais ou inferiores a 30 anos* que mais cumprem a política de acesso aberto à produção científica do ISCTE-IUL. Não obstante as diferenças encontradas apenas são estatisticamente relevantes as que dizem respeito à *Carreira Profissional* ($M-WZ=-2,207$; $p=0,027$) e ao *Departamento* ($KW=4,888$; $p=0,087$).

Curiosamente se fosse aprovado um mandato no ISCTE-IUL os valores mais elevados subiam e os mais baixos desciam, isto é, 82,5% *Cumpriria integralmente e concordaria*, apenas 7,7% *Cumpriria parcialmente*. De referir que nestas circunstâncias a percentagem dos que não cumpririam de todo é inexistente.

Quando distribuídos estes valores pelas diferentes variáveis de caracterização verificamos que, mais uma vez, são os *Investigadores*, pertencentes a *Departamentos e de Unidades de Investigação integradas na ESPP* e com *idades iguais ou superiores a 51 anos* que revelaram uma maior tendência para um maior nível de cumprimento caso o ISCTE-IUL aprovasse uma política obrigatória (mandato) de acesso aberto à produção científica dos seus afiliados. Contudo, apesar de algumas diferenças encontradas nenhuma delas revelou evidência estatística.

Quando questionados sobre *qual seria, de acordo com a sua opinião, a forma mais fácil de satisfazer os requisitos de uma política obrigatória (mandato) de acesso aberto à produção científica* um pouco mais de metade referiu a opção que *combina a publicação em revistas de acesso aberto e o depósito de artigos/documentos num repositório institucional/temático* (53,1%), sendo a segunda opção com uma percentagem mais alta (24,5%) *depositar uma cópia do artigo/documento num repositório institucional/temático*.

Se distribuirmos estes valores pelas diferentes variáveis de caracterização é possível apurar que qualquer uma das opções é igualmente valorizada quer por *Docentes* quer por *Investigadores*, apenas a opção *Não sei / Sem opinião* obteve uma percentagem maior por parte dos *Docentes*.

Mas, por exemplo, em termos de pertença Departamental a percentagem mais elevada surge na opção *Combinar a publicação em revistas de acesso aberto e o depósito de artigos/documentos num repositório institucional/temático* por parte dos docentes de *Departamentos pertencentes à ECSH*. Esta opção é igualmente a escolhida por parte dos pertencentes a *Unidades de Investigação* integradas na *ETA / EG*. Já no que toca ao *Sexo* as *Mulheres* seguem a tendência em termos globais pois a maior percentagem surge na opção *Combinar a publicação em revistas de acesso aberto e o depósito de artigos/documentos num repositório institucional/temático*. Na *Idade* também foi esta a opção escolhida por aqueles com idade igual ou superior a 51 anos.

Quando questionados sobre a *definição de embargos por parte dos editores e sobre quais os períodos de tempo a considerar para o efeito* a opção mais escolhida foi precisamente a de *acesso livre imediato* (37,8%). Com valores idênticos surge a opção *6 meses* e *Não sei / Sem opinião*, ambas com 23,8%.

Em termos de distribuição destes valores pelas variáveis de caracterização não se verificam diferenças acentuadas relativamente aos valores globais, ou seja, a maior parte dos respondentes estão distribuídos maioritariamente pela primeira opção e os restantes pela opção *6 meses* ou então não sabem ou não têm qualquer opinião.

Quando questionados sobre a *opinião relativamente ao princípio da disponibilização em acesso aberto dos dados científicos resultantes de projetos financiados por programas públicos* tomando em consideração que o acesso aberto pode ser relevante não apenas para as publicações mas também para os dados científicos pouco mais de 50% *Concorda sem restrições* face a 11, 2% que *Não concorda* de todo.

Uma percentagem considerável (35,0%) afirmou que *Concorda dependendo das condições*. Nesta opção foi dada a possibilidade dos respondentes indicarem quais seriam essas condições. Assim, foi possível agrupar as condições apontadas em três grandes categorias: 1) *Confidencialidade dos dados / anonimato*; 2) *Restrição do acesso / mediante autorização do coordenador do projeto*; 3) *Haver garantia de que todas as publicações que pudessem resultar do projeto tenham já sido produzidas, publicadas e divulgadas*.

Cruzando estes valores pelas variáveis de caracterização e no que respeita à *Carreira Profissional* não se verificam grandes diferenças em qualquer uma das categorias. Relativamente ao *Departamento* destacam-se os respondentes que pertencem à *ETA/EG* em qualquer uma das opções existentes, com 64,7% a *Concordarem sem restrições*, apenas 23,5% *Concordam dependendo das condições* e 11,8% a *Não concordarem*. No que respeita às condições apresentadas prendem-se essencialmente com questões de confidencialidade e proteção de dados privados.

No que respeita à *pertença a determinada Unidade de Investigação* ressaltam os membros daquelas que pertencem à ESPP pois 66,7% *Concordam sem restrições*, 23,8% *Concordam dependendo das condições* e 9,5% *Não concordam*. De referir que nesta ultima opção os membros da ETA/EG atingiram uma percentagem de 20% o que está em consonância com a variável anterior que dizia respeito ao departamento.

No que respeita às condições são referidas questões de confidencialidade e anonimato, salvaguarda de acordos com as instituições fornecedoras dos dados e assegurar a forma como os dados seriam publicados e quem os poria utilizar.

No que respeita ao *Sexo* e na opção *Concordo sem restrições* não se verificam grandes diferenças entre os dois grupos, mas na opção *Não concordo* as Mulheres apresentam uma percentagem de 15,4% face a 3,8 por parte dos Homens.

No tocante à *Idade* as diferenças entre os grupos não são muito acentuadas, destacando-se na opção *Não concordo* a faixa etária entre os 41 e os 50 com uma percentagem mais baixa do que a apurada nos outros grupos.

Foi, igualmente, possível apurar que quer a *Participação em projetos financiados no âmbito do 7º Programa-Quadro da Comissão Europeia* ($M-WZ=-2,158$; $p=0,031$) quer a *Participação noutros projetos* influencia a opinião dos participantes neste estudo relativamente ao princípio da disponibilização em acesso aberto dos dados científicos resultantes de projetos financiado por programas públicos ($M-WZ=-2,360$; $p=0,018$), ou seja, as diferenças apuradas têm relevância estatística.

Tabela 10 Opinião sobre o princípio da disponibilização em acesso aberto dos dados científicos resultantes de projetos financiado por programas públicos segundo as variáveis de caracterização

		Concordo sem restrições		Concordo dependendo das condições		Não concordo	
		n	%	n	%	n	%
Carreira Profissional	Docente	26	52,0	19	38,0	5	10,0
	Investigador	51	54,8	31	33,3	11	11,8
Deptº / Esc	ECSH	7	50,0	6	42,9	1	7,1
	ETA / EG	11	64,7	4	23,5	2	11,8
	ESPP	8	47,1	8	47,1	1	5,9
Unid / Esc	ECSH	20	44,4	19	42,2	6	13,3
	ETA / EG	2	40,0	2	40,0	1	20,0
	ESPP	28	66,7	10	23,8	4	9,5
Sexo	Feminino	48	52,7	29	31,9	14	15,4
	Masculino	29	55,8	21	40,4	2	3,8
Idade	≤ 30	14	48,3	12	41,4	3	10,3
	31-40	32	57,1	17	30,4	7	12,5
	41-50	22	56,4	14	35,9	3	7,7
	≥ 51	9	47,4	7	36,8	3	15,8

Carreira Profissional ($M-WZ=-0,185$; $p=0,854$) / Departamento/Escola ($KW= 0,655$; $p=0,721$) / Unidade de Investigação / Escola ($KW= 4,181$; $p=0,124$) / Sexo ($M-WZ=-0,889$; $p=0,374$) / Idade ($KW= 0,966$; $p=0,802$)

Sobre a Prática de Acesso Livre

Esta seção do questionário procurou obter informação sobre a prática efetiva do Acesso Aberto apurando se os participantes neste estudo já tinham disponibilizado publicações em acesso aberto e identificando igualmente os motivos pelos quais o fizeram ou não.

Quando questionados sobre o *número de publicações que melhor define a sua produção científica* a grande maioria (76,9%) indicou que tem entre 1 a 5 publicações.

Quando analisamos estes resultados de acordo com as variáveis de caracterização e no que respeita à *Carreira Profissional académica* percebemos que são os *Investigadores* que, de forma geral, menos produzem, pois embora com *Menos de 1 publicação por ano* apresentem uma percentagem superior (24,7%) à dos *Docentes* (10,0%), no escalão de *1 a 5 publicações por ano* são os *Docentes* que apresentam mais produção com 86% face aos 72,0% dos *Investigadores*. As diferenças encontradas entre estes dois grupos são estatisticamente significativas significando isso que a sua principal atividade profissional influencia a número de obras publicadas ($M-WZ=-1,992$; $p=0,046$).

Quanto à pertença *Departamental*, a grande maioria tem entre 1 a 5 publicações por ano em qualquer um dos grupos, com exceção dos membros da ETA/EG que apresentam 70,6% e que igualmente são os únicos que têm *Menos de 1 por ano* (29,4%). Estas diferenças embora não muito acentuadas são estatisticamente significativas, ou seja, a pertença departamental influencia o volume da produção científica dos seus membros ($KW= 10,126$; $p=0,006$).

Quanto à pertença a uma determinada *Unidade de Investigação* os valores apurados não demonstram grandes diferenças entre os grupos. De destacar apenas que com *Menos de 1 publicação por ano* surge com maior percentagem (28,9%) os membros pertencentes a Unidades integradas na ECSH e de 1 a 5 publicações por ano surge com maior percentagem os membros pertencentes a unidades integradas na ETA / EG.

Quanto ao *Sexo* as diferenças apuradas são mínimas e no que refere à *Idade* destacam-se os respondentes com idade igual ou superior a 51 anos com maior percentagem (89,8%) na opção *De 1 a 5 publicações por ano*.

É de salientar que, quer a participação em projetos financiados no âmbito do 7º PQ da CE ($M-WZ=-3,050$; $p=0,002$), quer a participação noutros projetos ($M-WZ=-3,138$; $p=0,002$) condicionam o volume da produção científica dos docentes e investigadores.

39,2% dos respondentes afirmam ter disponibilizado as suas publicações no Repositório do ISCTE-IUL, 37,1% em Plataforma de partilha de conteúdos científicos (Academia.edu, ...) e 35,7% através da publicação numa revista em acesso aberto. De salientar, que do total dos respondentes apenas 18,9% afirmaram não ter disponibilizado as suas publicações em acesso aberto.

No que respeita à *Carreira Profissional* não se verificam diferenças assinaláveis em nenhuma das categorias.

Já no que respeita ao *Departamento* é possível verificar algumas diferenças mais acentuadas, nomeadamente, na *disponibilização de publicações em acesso aberto através do depósito no repositório do ISCTE-IUL* (47,6%) ou através da *publicação numa revista em acesso aberto* (55,6%) por parte dos membros dos departamentos pertencentes à ESPP; na *disponibilização de publicações em acesso aberto através do depósito num repositório em acesso aberto* (63,6%) ou através do *depósito num repositório em acesso aberto e da publicação numa revista em acesso aberto* (50,0%) por parte dos membros dos departamentos pertencentes à ETA ou EG. Destas relações apenas as encontradas nesta variável e a disponibilização através do repositório do ISCTE-IUL ($X^2=4,616$; $p=0,099$), de numa revista em acesso livre ($X^2=5,190$; $p=0,075$) e através do depósito num repositório em acesso aberto ($X^2=5,448$; $p=0,066$) revela evidência estatística.

No que respeita à *pertença a determinada unidade de investigação* é de assinalar que no caso das que pertencem à ETA ou EG a maior percentagem surge na última opção que assinala a não disponibilização de publicações em acesso livre. Por outro lado, a publicação através do Repositório do ISCTE-IUL ou de repositório em acesso livre surge da parte de membros das Unidades de Investigação pertencentes à ESPP. Já as opções que assinalam a publicação numa *revista em acesso aberto, através do depósito num repositório em acesso aberto e da publicação numa revista em acesso aberto* ou numa *plataforma de partilha de conteúdos científicos (Academia.edu, ...)* obtêm percentagens mais elevadas de membros das Unidades de Investigação pertencentes à ECSH.

Relativamente ao *Sexo* podemos assinalar algumas diferenças significativas nomeadamente no que respeita à publicação através da publicação numa revista em acesso aberto ($X^2=2,721$; $p=0,099$) e à através do depósito num repositório em acesso aberto ($X^2=4,455$; $p=0,035$).

No tocante à *Idade* é possível apurar que os respondentes com idade compreendida entre os 31 e os 40 anos são os que têm percentagens mais elevadas em qualquer uma das possibilidades de publicação apresentadas. Os participantes neste estudo com idades compreendidas entre os 41 e os 50 anos foram os que maior percentagem obtiveram na opção *Não*. Da relação da *Idade* com cada uma das opções de publicação apenas as diferenças encontradas nos respondentes que publicam numa plataforma de partilha de conteúdos científicos (Academia.edu, ...) se revelaram estatisticamente significativas ($X^2=11,717$; $p=0,008$).

Tabela 11 Disponibilização de publicações em acesso aberto segundo as variáveis de caracterização

		Sim, através do depósito no repositório do ISCTE-IUL		Sim, através do depósito num repositório em acesso aberto		Sim, através da publicação numa revista em acesso aberto		Sim, através do depósito num repositório em acesso aberto e da publicação numa revista em acesso aberto		Sim, numa plataforma de partilha de conteúdos científicos (Academia.edu, ...)		Não	
		n	%	n	%	n	%	n	%	n	%	n	%
Carreira Profissional	Docente	21	37,5	11	39,3	18	35,3	6	28,6	16	30,2	11	40,7
	Investigador	35	62,5	17	60,7	33	64,7	15	71,4	37	69,8	16	59,3
Deptº / Esc	ECSH	7	33,3	1	9,1	4	22,2	1	16,7	3	20,0	3	30,0
	ETA ou EG	4	19,0	7	63,6	4	22,2	3	50,0	6	40,0	5	50,0
	ESPP	10	47,6	3	27,3	10	55,6	2	33,3	6	40,0	2	20,0
Unid / Esc	ECSH	12	35,3	6	37,5	19	57,6	8	53,3	19	51,4	10	62,5
	ETA ou EG	2	5,9	0	,0	2	6,1	1	6,7	1	2,7	2	12,5
	ESPP	20	58,8	10	62,5	12	36,4	6	40,0	17	45,9	4	25,0
Sexo	Feminino	38	67,9	13	46,4	37	72,5	13	61,9	36	67,9	16	59,3
	Masculino	18	32,1	15	53,6	14	27,5	8	38,1	17	32,1	11	40,7
Idade	≤ 30	11	19,6	2	7,1	12	23,5	3	14,3	11	20,8	6	22,2
	31-40	22	39,3	15	53,6	18	35,3	9	42,9	27	50,9	7	25,9
	41-50	12	21,4	4	14,3	13	25,5	6	28,6	6	11,3	13	48,1
	≥ 51	11	19,6	7	25,0	8	15,7	3	14,3	9	17,0	1	3,7

Quando inquiridos sobre o número aproximado de *artigos em acesso aberto depositados no Repositório do ISCTE-IUL*, 46,6% depositaram entre 1 a 9 artigos, 1,7% entre 10 a 19 artigos e 4,3% 20 ou mais artigos.

Relativamente ao número de *artigos em acesso aberto depositados em repositórios institucionais*, 35,3% depositaram entre 1 a 9 artigos, 1,7% entre 10 a 19 artigos e igual percentagem tem 20 ou mais artigos.

Relativamente ao número de *artigos em acesso aberto depositados em repositórios temáticos*, 24,3% depositaram entre 1 a 9 artigos, 2,6% entre 10 a 19 artigos e apenas 0,9 % tem 20 ou mais artigos depositados em repositórios temáticos.

No que respeita ao número de *artigos em revistas científicas em acesso aberto*, 57,4% depositaram entre 1 a 9 artigos e 1,7% entre 10 a 19 artigos.

Foi igualmente importante tentar perceber *quais os motivos que levavam os Docentes e Investigadores a depositarem as suas publicações científicas em repositórios institucionais/temáticos* e apuramos que os dois principais motivos são precisamente aqueles que mais utilizamos para promoção da “utilidade” de um repositório, isto é, o *Aumento da visibilidade e o impacto (citações) da investigação* (86,0%) e a *Rapidez de disseminação dos resultados da investigação* (80,4%). Como motivos menos valorizados surge o facto de *Poder delegar esta tarefa em outra pessoa* (2,1%) e *Haver um requisito do financiador da investigação* (4,9%). Os resultados obtidos vão ao encontro dos resultados obtidos em estudos realizados por Swan (2011) que evidenciam que maior visibilidade se traduz no aumento d número de citações.

Como *motivos para não depositar as suas publicações científicas num repositório institucional/temático* surge como principal a *Falta de conhecimento sobre as políticas de depósito dos editores* (32,2%), logo seguido do *Receio de violar os direitos de autor* com 31,5% e da *Falta de conhecimentos técnicos e formação no uso do sistema* (30,8%).

Os motivos menos seleccionados foram o *Os meus colegas também não o fazem* (2,1%) e *Não reconheço vantagens no acesso aberto* (2,8%).

No final do questionário era pedido aos respondentes que formulassem comentários no âmbito da temática em estudo. Foram recebidos 19 comentários. A análise desta informação passou pela categorização de cada resposta, sendo que, nalguns casos, um único comentário se enquadrava em mais do que uma categoria. Foram excluídos 3 comentários por não terem relevância para o estudo.

As categorias definidas foram as seguintes:

Acessibilidade → É referida a falta de canais de difusão da produção científica em modelos facilmente acessíveis pela população em geral. Alguns participantes mencionam as páginas pessoais como locais preferenciais para a disponibilização da sua produção científica.

Avaliação / Impacto → Nesta categoria é salientado o facto de nos processos de avaliação serem valorizadas as publicações em revistas com fatores de impacto elevado que não são compatíveis com o acesso aberto. As instituições de acolhimento não podem exigir publicação absoluta em acesso aberto e depois baixar o *ranking* do investigador por não publicar em revistas de referência.

Direitos de autor → São referidas questões como o receio de violar as políticas de direitos de autor editoras.

Modelos de publicação → Verifica-se um nível elevado de concordância com a disponibilização em acesso livre desde que não sejam imputados custos aos autores. É referida a necessidade de conciliar as vantagens da publicação em revistas (revisão por pares, classificação, entre outras) com as vantagens do acesso aberto.

Revisão por pares → É mencionada a necessidade de garantir a continuação da existência de *peer review* mesmo em publicações em acesso aberto de forma a que a qualidade das mesmas esteja assegurada. Este aspeto está diretamente relacionado com a dimensão da confiabilidade “que decorre de todos os processos que asseguram todos os processos que asseguram a qualidade dos documentos, nomeadamente, o da revisão por pares assim como o facto de os repositórios serem plataformas que permitem o acesso a materiais que, na sua maioria, já foram publicados ou, de alguma



forma, avaliados, como é o caso das teses de doutoramento e das dissertações de mestrado.” (Amante, 2013, 196). Contudo, a avaliação da qualidade científica de um texto pode ser aferida através da utilização de outros parâmetros, por exemplo, o prestígio das revistas, a reputação dos autores dos trabalhos e das instituições a que pertencem (Cronin, 2005).

Através desta análise tornou-se possível verificar que os investigadores e docentes do ISCTE-IUL têm perceções e concepções divergentes relativamente ao acesso aberto, às formas de o providenciar, aos direitos de autor, às licenças e permissões das revistas onde publicam artigos, evidenciando desconhecimento da realidade. Do conjunto das preocupações, destacam-se as associadas aos custos do acesso aberto e às suas implicações na avaliação dos investigadores e docentes e das suas publicações.

CONCLUSÕES E NOTAS FINAIS

O presente estudo, através de inquérito por questionário, permitiu recolher informação sobre o nível de conhecimento e as perceções dos investigadores e docentes do ISCTE-IUL relativamente ao acesso aberto à informação científica.

Foi, igualmente, possível apurar que o nível de conhecimento, quer da política europeia, quer da política institucional é bastante reduzido. No ISCTE-IUL e, dos participantes que têm conhecimento da política em vigor, o nível de cumprimento é acima da média com tendência a aumentar caso um mandato fosse aprovado. De referir que este conhecimento é de alguma forma influenciado pela participação em projetos de investigação, nomeadamente, aquelas financiados no âmbito do 7º PQ da CE.

Os resultados permitiram, igualmente, verificar que cerca de 110 investigadores/docentes têm entre 1 a 5 publicações por ano. Destes, apenas uma percentagem de 18,9% não tem nenhuma em acesso livre. As razões apontadas são maioritariamente a falta de conhecimento sobre as políticas de depósito das editoras e conseqüente receio de violar os direitos de autor, a falta de conhecimentos técnicos e formação no uso do sistema, a falta de tempo para realizar o depósito e o desconhecimento sobre o carácter obrigatório das políticas institucionais.

Assim, dos resultados obtidos foi possível retirar duas conclusões:

→ Existe espaço para a definição de uma política mandatária, ou seja, que exija a disponibilização em acesso aberto das publicações dos investigadores e docentes, uma vez que estes evidenciam um elevado grau de concordância com a mesma.



→ Foi igualmente possível apurar que se torna necessário um reforço no que respeita à realização de ações de divulgação e sensibilização sobre o acesso aberto sobre questões relacionadas com os direitos de autor bem como sobre as políticas definidas atualmente. Com este objetivo “iniciámos recentemente um processo de contacto individualizado com docentes e com investigadores solicitando a sua colaboração no depósito de artigos de sua autoria identificados através da *Web of Science* que não está no Repositório do ISCTE-IUL. Trata-se quer de artigos que já se encontram sem período de embargo quer de artigos mais recentes na sua versão *preprint* ou *postprint*, relativamente aos quais foram por nós confirmados os direitos e períodos de embargo das revistas onde foram publicados no *site* do Sherpa/Romeo.” (Amante, 2013, 199)

→ Também a criação de procedimentos e ferramentas que diminuam a duplicação de tarefas se torna premente.

→ Por último, devido ao contexto de crise e de decorrentes restrições financeiras que obriga a uma análise muito ponderada dos custos envolvidos na manutenção dos Repositórios (infraestrutura tecnológica necessária à sua implementação e manutenção, procedimentos envolvidos no fornecimento de serviços de valor acrescentado, estratégias de preservação) parece-nos inevitável a adoção de indicadores de desempenho que permitam avaliar o sucesso dos mesmos. (Cassella, 2010, 211)

BIBLIOGRAFIA

Amante, Maria João (2013). Acesso Aberto@ISCTE-IUL. In: *Uma década de Acesso Aberto na UMinho e no Mundo* (eds Eloy Rodrigues, Alma Swan e Ana Alice Baptista), pp. 187-202.

Cassella, Maria (2010). Institutional repositories: an internal and external perspective on the value of IRs for researchers' communities. *Liber Quarterly*, vol. 20, nº 2, pp. 210-225.

Cronin, Blaise (2005). *The hand of science: academic writing and its rewards*. Oxford: Scarecrow Press Inc. ISBN 0-8108-5292-9. 214 p.

Swan, Alma (2011). Institutional repositories – now and next. In: *University Libraries and Digital Learning Environments* (eds Penny Dale, Jill Beard e Matt Holland), Ashgate Publishing. <http://eprints.soton.ac.uk/id/eprint/271471>.



SERVICIOS DE APOYO A LA GESTIÓN DE PUBLICACIONES DE ACCESO ABIERTO: EL CASO DE LA UNIVERSIDAD DE COSTA RICA

Meilyn Garro-Acón, Jorge Polanco-Cortés, Johan Espinoza Rojas

RESUMEN

Este trabajo describe la experiencia del equipo UCRIndex (Latindex Costa Rica) en la implementación de un modelo de servicios de apoyo a la gestión de publicaciones aplicado a revistas de acceso abierto de la Universidad de Costa Rica. Desde hace varios años, la Universidad apoya la gestión editorial de sus publicaciones periódicas con el fin de mejorar su calidad, visibilidad e impacto. El modelo parte de que los encargados de publicaciones académicas deben adquirir una serie de capacidades y conocimientos técnicos para lograr visibilidad y mejor posicionamiento en los distintos espacios académicos, virtuales o físicos. La presente sistematización plantea que dichas características no siempre son fáciles de alcanzar por los editores académicos, siendo necesario generar una serie de actividades y procesos que permitan la evolución y el perfeccionamiento de las publicaciones de forma que se aprovechen los esfuerzos al aplicar las mejoras a la mayor cantidad de revistas posible. La aplicación de este modelo de apoyo ha resultado en una completa plataforma de servicios para revistas que incluye la creación de propuestas para la actualización y capacitación de los editores, investigación sobre tecnologías y otros temas, gestión de servicios comunes (contratación con sistemas de identificadores únicos, actualización de contenidos en web, marcaje de artículos para servicios de indexación, creación de respaldos y preservación de contenido), seguimiento y evaluación de algunos procesos, así como la creación de espacios de socialización de experiencias para las revistas. Los resultados de este trabajo se reflejan en el aumento en la cantidad de revistas con sus números completos en línea, aumento en la cantidad de índices selectivos a los cuales pertenecen las revistas, mayor y mejor uso de herramientas en línea por parte de los editores, mejoramiento y automatización del proceso de gestión editorial, y mejoramiento de la calidad formal de las publicaciones.

Palabras clave: acceso abierto; publicación científica; transferencia de conocimientos; gestión del conocimiento



ABSTRACT

This paper describes the experience of the UCRIndex team (Latindex Costa Rica) in the implementation of a model of support service management applied to open access publications at the University of Costa Rica. For several years, the University had support the editorial management of its journals in order to improve its quality, visibility and impact. This model suggest that academic editors should develop several technical skills in order to achieve exposure in academic spaces, either physically or virtually. These characteristics will not always be easy to acquire for some editors. For this reason, it is necessary to implement some activides and training in order to achieve refinement and improve the development of many magazines as possible. This model has led to a service platform for journals, which include: proposals for updates and guidelines for the editors, technology research and other issues, common services (systems of unique identifiers, web content updating, marking articles for indexing services, backups and preservation), monitoring and evaluating some processes, and meetings for socializing and sharing experiences of the different journals. The results of this work are reflected in the increasing number of journals with complete online collection and the indexes where they belong, the increasing of the use of online tools by the publishers, improving and automation the editorial management, and formal quality of publications.

Keywords: Open access; scientific publications; know-how transfer; knowledge management

INTRODUCCIÓN

Las revistas académicas son un medio reconocido para la divulgación de la investigación, gracias a la evaluación por pares y la aprobación del consejo editorial que validan la calidad de la investigación. En Costa Rica, el mantenimiento de las revistas científicas es financiado por instituciones estatales, principalmente universidades, que con frecuencia asumen costos de edición elevados debido a la gran cantidad de revistas que puede llegar a editar una misma institución. El caso presentado en esta ponencia se basa en la experiencia de la Universidad de Costa Rica (UCR), la cual ha editado al menos 66 revistas académicas a lo largo de su historia, cada una financiada por diferentes dependencias universitarias, sin coordinar dicha edición e impulsadas por intereses de particulares. Esta situación, generó una falta de control sobre los fondos asignados a cada publicación y generó varias revistas con una misma cobertura temática en las que, muchas veces, no se aseguraba personal y tiempo para mantener la gestión de cada publicación en el tiempo.



Para normalizar esta situación, la UCR comenzó un proceso de revisión de los recursos asignados a las publicaciones, el estado de sus indicadores de calidad, una moratoria en la creación de revistas con la misma cobertura temática, y la implementación de una serie de servicios para la gestión de publicaciones académicas de acceso abierto. Instituciones como Universidad Nacional de la Plata en Argentina, Universidad de Valencia¹ han desarrollado servicios de apoyo a la gestión editorial con el fin de que sus publicaciones brinden mejores servicios y respondan ante las necesidades tecnológicas actuales. Estos servicios se dedican al mantenimiento, actualización y desarrollo de sistemas para revistas, repositorios, congresos y bibliotecas, entre otros. En el caso de la UCR el servicio de apoyo se encuentra en la Vicerrectoría de Investigación, a cargo del equipo UCRIndex se encarga de brindar capacitación, actualización profesional, investigación y acompañamiento a los editores de sus revistas, así como mantenimiento, actualización y adaptación del sistema Open Journal Systems (OJS)², para el portal de revistas, y DSPACE, para repositorios. Al conformar el equipo UCRIndex, se esperaba ofrecer un mejor servicio técnico a las revistas, realizar una reducción de costos de operación, controlar la inversión en cada una de las publicaciones y monitorear indicadores de calidad (por medio de herramientas generadas por Latindex Costa Rica y el índice UCRIndex). Con ello, se espera que revistas mejoraran su calidad y visibilidad, además de realizar una evaluación continua de las publicaciones.

ANTECEDENTES

Como parte del esfuerzo para mejorar la calidad de las revistas institucionales, la UCR se convirtió en la institución responsable de Latindex a nivel nacional a partir del año 2002, siendo la primera institución a nivel centroamericano en asociarse al sistema. Córdoba (2006) menciona que en el año 2002 el catálogo Latindex incluía 11 revistas, pero para el año 2006 había aumentado a 26 gracias a las capacitaciones y trabajo con los grupos editoriales. Según datos de Latindex, para julio de 2014 hay 105 revistas en el catálogo a nivel nacional. Se estima que este aumento en la cantidad de títulos se ha dado, en gran medida, gracias a la capacitación y concientización que el equipo de Latindex Costa Rica comunica en sus actividades.

¹Página de Servicios de Biblioteca y Documentación de la Universidad de Valencia
<http://www.uv.es/uvweb/servei-biblioteques-documentacio/es/serveis/suport-investigacio-edicio-cientifica/suport-edicio-cientifica/publicar-una-revista-1285872199666.html>

² Sistema administrador de contenido libre especial para la publicación de revistas.



Con el tiempo, la Vicerrectoría de Investigación de la UCR basada en la idea de una efectiva gestión de calidad y la mejora continua de sus publicaciones, aumentó y formalizó a los colaboradores de Latindex Costa Rica en el 2012, convirtiendo al equipo UCRIndex en un grupo de apoyo a las publicaciones académicas. Se creó así el índice UCRIndex, que categoriza a las revistas en cuartiles según unos criterios que buscan aumentar la calidad y exigencias de la Universidad con las revistas. Este índice UCRIndex³ tomó los criterios Latindex como parte de sus indicadores, lo que motivó a los editores a cumplir a cabalidad las evaluaciones Latindex.

Además, se le brindó este equipo de apoyo la confianza para liderar la divulgación científica y académica desde la Vicerrectoría de Investigación. Parte del trabajo del equipo es el posicionamiento en la web de la producción académica mediante la estandarización y mantenimientos de las herramientas de divulgación científica como las revistas y los documentos académicos, lo que brindó a los editores de revistas una herramienta que automatiza parte del proceso de gestión editorial, permite la publicación automática de los fascículos y le facilita a los investigadores y estudiantes una herramienta para la divulgación automática de las publicaciones producidas desde la Universidad.

Poco a poco se ha desarrollado una metodología para la elección de servicios necesarios para las publicaciones académicas y una serie de servicios brindados por el equipo UCRIndex. Este es un proceso que lleva aproximadamente 10 años y muestra resultados positivos desde temprano en su implementación.

EVOLUCIÓN DE LOS SERVICIOS DE APOYO PARA LA GESTIÓN DE PUBLICACIONES

Los servicios de apoyo comenzaron como una respuesta a la necesidad de profesionalización de la labor editorial. En algunos casos, quienes se encargan de las revistas incluso, desconocían la aplicación de criterios básicos de calidad editorial y no tenían conocimiento de las tareas básicas para editar una revista. Otros, surgieron debido a necesidades de apoyo técnico y de actualización de conocimientos de los editores a nivel tecnológico. Por ejemplo, después de algunas pruebas con implementación de software para gestión de revistas por parte de los editores, se concluyó que es más efectivo y eficiente centralizar ciertas tareas de instalación y mantenimiento de software. Estas condiciones se evidenciaban en los problemas mencionados anteriormente y repercutían directamente en la visibilidad y calidad de las publicaciones académicas de la Universidad.

³ Para ver la metodología del índice UCRIndex visite www.ucrindex.ucr.ac.cr



Con la implementación de soluciones para algunos problemas, se reflejaron nuevas oportunidades de mejora que pueden ser resueltas utilizando el mismo esquema de trabajo: un servicio común brindado por un equipo, que puede ser aprovechado por un grupo de publicaciones sin importar la temática. El grupo de beneficiarios puede ser extendido a tantas personas como instituciones participen, y en el equipo de trabajo pueden participar diferentes instituciones.

Este mismo modelo de servicios se encuentra en etapa inicial de aplicación en repositorios institucionales de las universidades en el Consejo Nacional de Rectores (CONARE), ente coordinador de las universidades públicas. Se busca con esto, la inclusión de procesos académicos como evaluación de la producción científica y entrega de informes de centros de investigación, entrega de documentos para evaluación de régimen académico, acopio de documentos para acreditación de carreras, entre otros; para más tarde hacer medición de la ciencia por medio de los portales de revistas y repositorios institucionales.

A continuación se detallan algunos de los servicios implementados a nivel de revistas, primer campo de aplicación de este modelo de servicios por parte de este equipo de trabajo de la UCR.

1. Evaluación de criterios de calidad de revistas.
2. Capacitación para editores académicos y técnicos.
3. Plataforma de divulgación web para publicaciones académicas.
4. Apoyo ingreso a índices.
5. Análisis de ventajas y oportunidades para revistas.
6. Desarrollo de software para el mejoramiento común.
7. Preservación digital.

Evaluación de criterios de calidad de revistas: Fue uno de los primeros servicios ofrecidos. Permite identificar necesidades y oportunidades de mejora a nivel de calidad editorial, así como monitorear la evolución de las revistas por medio de indicadores. El primer sistema de evaluación implementado fue Latindex (aplicado desde 2002). Más adelante se vio la necesidad de incrementar los indicadores monitoreados, por lo que se desarrolló el índice UCRIndex⁴, un índice institucional cuyo objetivo es potenciar el aumento en la calidad de las revistas mediante el cumplimiento de indicadores que les serán solicitados para ingresar en índices internacionales.

En este índice se identifican las carencias de las revistas para implementar acciones que ayuden a los editores a calificar positivamente en menos tiempo, permitiendo ofrecer ayuda técnica a nivel informático y legal. Los indicadores incluidos en el índice UCRIndex se definen dependiendo de las tendencias y exigencias de índices internacionales que indexen publicaciones de acceso abierto.

⁴Basado en la experiencia aportada por sistemas como Publindex-Colombia, el Índice de Revistas Mexicanas, el Núcleo Básico de Revistas Argentinas, Qualis de Brasil y DICE de España, así como el conocimiento de la realidad de las revistas costarricenses.



La evaluación Latindex y UCRIndex son obligatorias para todas las revistas académicas de la Universidad de Costa Rica y el mejoramiento en los indicadores puede ser tomado en cuenta por la autoridades para tomar decisiones sobre la asignación de recursos económicos. Las evaluaciones Latindex se aplican por solicitud al resto de revistas nacionales, pero dados los buenos resultados con el índice UCRIndex, se está considerando aplicar un índice similar a todas las revistas de las universidades públicas por medio de un proyecto interuniversitario.

Capacitación para editores académicos y técnicos: Una vez identificadas las necesidades a nivel de calidad editorial, se define cuáles pueden ser resueltas con capacitación y actualización técnica de los editores y otros miembros del personal de las revistas. Con base en esos datos, se desarrolla un calendario de actividades y capacitación anual que incluye siempre la formación de nuevos editores (debido a la rotación de personal o relevo generacional, y a la necesidad de que otros participantes del proceso editorial conozcan los principios y generalidades de una revista). Este programa de capacitación permite la solicitud de contenidos y el planteamiento de problemas por parte de los editores y hasta ahora ha incluido temas como gestión editorial para nuevos editores, software para gestión de revistas, definición de políticas editoriales y derechos de autor, aplicación de normas de citación, cumplimiento de criterios para ingresar a índices, normas para la creación de resúmenes y palabras clave, entre otros temas.

Este servicio de capacitación existe desde 2007. Desde entonces, las sesiones de capacitación se han convertido en espacios para la socialización de experiencias entre responsables de revistas, haciendo posible una mayor comunicación y colaboración entre los mismos. No toda la formación de los editores se encuentra en manos de un mismo equipo facilitador, sino que se incentiva la participación activa de los editores para encontrar problemas y soluciones comunes. Asimismo motivamos a la participación de especialistas internacionales que actualizan profesionalmente a los editores y al equipo de apoyo en temas de actualidad internacional.

Cabe resaltar que el programa de capacitación ya funciona a nivel interuniversitario y permite la organización de actividades por parte de las coordinaciones de cada institución participante. Además, la capacitación incluye la coordinación con otras instancias universitarias cuyos procesos pueden afectar a las revistas, tales como la migración a Software Libre.



Plataforma de divulgación web para publicaciones académicas: Aparte de las evaluaciones de revistas, inicialmente se identificó una carencia de visibilidad web para casi todas las revistas institucionales, aún más para aquellas que no contaban con una versión electrónica o un sitio web para divulgar sus contenidos. Motivados por esa preocupación, en 2007 se creó el primer repositorio de revistas de la UCR⁵, el cual describía y divulgaba cada uno de los artículos con miras a integrar la producción de la UCR en la web semántica, ofreciendo un espacio para que los usuarios ingresaran de manera gratuita. En esta etapa, sólo se prestaba el servicio de descripción y subida de documentos a ese repositorio.

Durante el funcionamiento de esta primera versión de portal de revistas se identificaron otros problemas como la falta mantenimiento y desactualización de contenidos en los servidores de los sitios de las revistas, así como la falta de tiempo de los editores para aprender y realizar algunas tareas que le corresponderían a un editor técnico o un web máster. A raíz de esto, se tomó la decisión ofrecer un servicio que resolviera algunos de esos problemas de manera compartida. Así comenzó el proyecto de implementación de un portal de revistas en Open Journal Systems (OJS) administrado por el equipo UCRIndex encargado del mantenimiento, actualización de software, capacitación para editores y soporte técnico. Adicionalmente, se integró el servicio de subida de artículos.

Actualmente, se trabaja en la capacitación de los editores para la utilización de más herramientas ofrecidas por el software y en la promoción de nuevos formatos para la divulgación de los artículos.

Apoyo ingreso a índices: Algunos índices demandan que sus revistas realicen trabajos de marcaje de artículos, definición de políticas y documentación de sus procesos editoriales. Por esta razón se implementó un servicio de apoyo para el marcaje de artículos y de asesoría legal para la definición de políticas. Al mismo tiempo, se mantiene comunicación con algunos índices selectivos y se invita a las revistas a cumplir sus criterios de manera anticipada.

El apoyo para ingresar a índices también incluye la generación de espacios para la socialización del conocimiento, la discusión de implicaciones del cumplimiento de algunos criterios y el apoyo con el marcaje de artículos en el caso de SciELO.

⁵El repositorio de revistas Latindex UCR estuvo en línea desde el 2007 hasta 2013, cuando su contenido fue completamente migrado al Portal de Revistas de la Universidad de Costa Rica. se puede ver una imagen del 2007 aquí <http://web.archive.org/web/20070502071706/http://www.latindex.ucr.ac.cr/>



Análisis de ventajas y oportunidades para revistas: Con la implementación de este modelo de servicios para publicaciones, el equipo de trabajo asumió la tarea de buscar nuevas oportunidades de mejora y analizar propuestas para determinar sus próximas acciones, así como la forma de coordinarlas con otros equipos participantes⁶. La línea general para buscar estas mejoras contempla la implementación de Software Libre, coordinar acciones con equipos de otras instituciones y buscar la manera de compartir inversiones con ellos para aumentar la cantidad de beneficiarios y disminuir la duplicación de tareas por institución.

Para formalizar esta cooperación interuniversitaria, se creó una subcomisión interinstitucional de CONARE⁷ en la que participan 4 universidades públicas con un presupuesto compartido para la ejecución de acciones orientadas al mejoramiento de sus publicaciones académicas periódicas a nivel de gestión editorial y visibilidad en Internet.

Desarrollo de software para el mejoramiento común: Gracias a la implementación de Software Libre, es posible desarrollar y mejorar herramientas necesarios para el fortalecimiento de las publicaciones. Esto incluye el mejoramiento y desarrollo de complementos para software de divulgación y gestión de publicaciones, así como la búsqueda de soluciones automáticas para la normalización de metadatos y la optimización de procesos que necesitan mucho recurso humano.

En este sentido, se adicionó el servicio de desarrollo de plugins al Portal de Revistas. El objetivo es coordinar con los otros portales de las universidades públicas para compartir los desarrollos de software e identificar nuevas necesidades de desarrollo para posteriormente proponer los aportes al Public Knowledge Project (PKP)⁸. En este momento se está trabajando en el desarrollo de un plugin para exportar, desde OJS, un archivo XML que cumpla con las nuevas especificaciones de SciELO, lo que reduciría costo y tiempo en el marcaje de artículos. Este plugin la posibilidad de hacer marcaje automático de las referencias mediante una mejora al plugin de citas que ya existe.

⁶Es ente caso, los otros equipos son los encargados de los portales de revistas de las otras universidades públicas de Costa Rica.

⁷Inicialmente, se creó un proyecto en la Comisión Nacional de Rectores (CONARE), ente encargado de la coordinación de acciones sistémicas entre las Instituciones de Educación Superior Universitaria Estatal en Costa Rica, el cual se convirtió en subcomisión en 2014 al reconocerse su importancia de sus acciones y su pertinencia en el tiempo para el mejoramiento de la divulgación de la investigación en las universidades.

⁸ Iniciativa encargada del desarrollo y mantenimiento del Open Journal Systems.



Preservación digital: Este es un servicio reciente cuya que está en proceso de coordinación de esfuerzos para implementar servicio de preservación digital latinoamericano para la producción científica y académica. La propuesta busca incluir portales de revistas, repositorios, catálogos de bibliotecas y archivos institucionales en una red cuyo desarrollo sea soportado por equipos internacionales encargados de esos sitios. Esta forma de trabajo tiende a la inclusión de pequeños equipos de desarrollo que se reparten el trabajo y comparten los beneficios.

Como parte de esta política de preservación se incluye la asignación de identificadores únicos (handle, DOI, por el momento) a los artículos de revistas académicas, respaldos y coordinación institucional para la aprobación de políticas de preservación que incluyan las necesidades de las publicaciones universitarias.

METODOLOGÍA DE TRABAJO

Al desarrollar este modelo de servicio institucional para promover la amplia difusión de los productos académicos se parte de que a) los editores de publicaciones académicas deben ajustar el proceso editorial a las nuevas formas de publicación y distribución del conocimiento; y b) es necesario gestionar servicios comunes para la promoción de la más amplia difusión de las producciones académicas. En este proceso de actualización, se toma en cuenta la necesidad de actualización técnica y la aplicación, cada vez más rigurosa, de estándares calidad de publicaciones científicas y espacios web para la divulgación académica.

La elección de los servicios que se prestan se realiza con base en datos obtenidos mediante:

- a) Identificación de necesidades por medio de contacto con editores y consejos editoriales. Este proceso se hace sistematizando observaciones presentadas en espacios para las socialización de experiencias.
- b) Identificación de carencias por medio de evaluación de revistas usando indicadores de calidad (Latindex, índice UCRIndex).
- c) Revisión de criterios de selección de índices y comparación con estado de las revistas.
- d) Monitoreo de rotación de personal de revistas para identificar y actualizar a las personas que necesitan formación inicial para mejorar su gestión en las publicaciones.
- e) Análisis de otras tendencias y recomendaciones internacionales relacionadas con divulgación científica.
- f) Investigación a nivel técnico (infraestructura, software, capacitación) y a nivel legal (conflictos y dudas comunes) para identificar oportunidades de mejora y necesidades de desarrollo.



Cada oportunidad de mejora identificada se somete a un análisis para determinar su viabilidad (inversión de tiempo y recursos económicos, beneficios esperados). Si la propuesta es viable, se solicita apoyo político e institucional para su implementación.

RESULTADOS

A partir de la implementación del modelo de servicios de apoyo a la gestión de publicaciones aplicado a las revistas de acceso abierto de la UCR se han logrado una serie de resultados que reflejan mejoras en la calidad de las publicaciones, aumento de la presencia en web y mayor ingreso en índices.

Algunos de los resultados más significativos son:

Disposición en línea de todas las revistas científicas de la UCR en acceso abierto

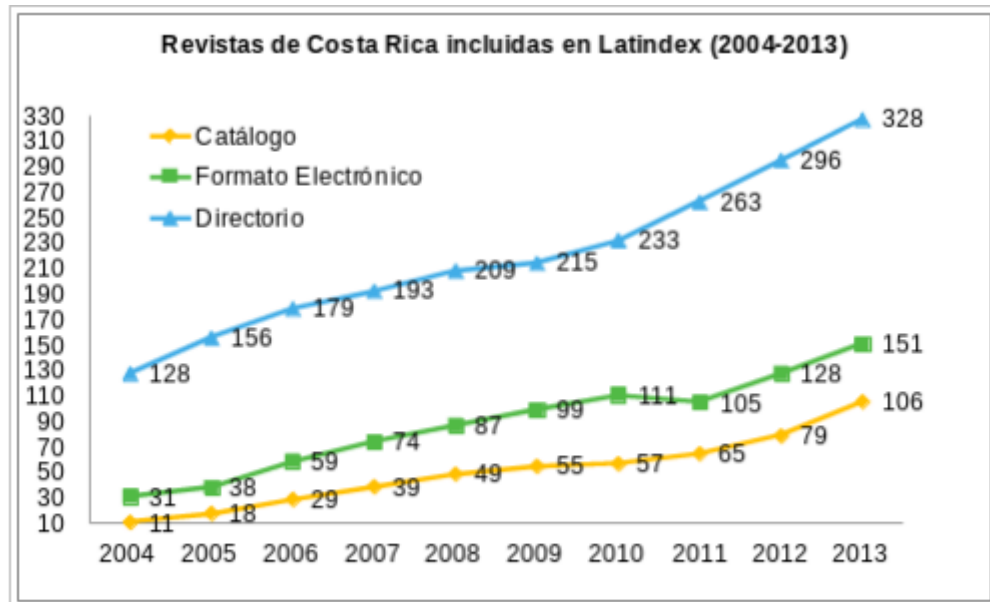
A pesar de que estas publicaciones siempre han sido de acceso abierto, muchas no contaban con el apoyo técnico para sostener sus propios sitios web y menos para subir todos sus números anteriores. El proceso de actualización de fascículos en línea comenzó en 2007 con 622 artículos en el antiguo repositorio Latindex UCR, para 2011 ya contaba con 4483 de 44 títulos, y a principios de 2014⁹ se contabilizaban más de 9500 artículos de las 52 revistas activas de la Universidad y continúa el proceso de subir los fascículos antiguos en el Portal de Revistas de la UCR (www.revistas.ucr.ac.cr). Antes de la implementación del Portal e Revistas, unas de 3 revistas contaban con sus propios sitios en OJS, pero fue hasta la creación del mismo que se inició un proceso de capacitación constante en el uso del software. El uso de este software mejoró la visibilidad web de los artículos, mediante el uso de metadatos que describen cada documento, algo que no se hacía cuando se subían los documentos en otros sistemas administradores de contenido (CMS).

Mejoramiento de la calidad editorial

Otro logro significativo es el aumento en el cumplimiento de criterios de calidad editorial. En cuanto a Latindex, como se puede observar el gráfico 1, desde el 2004 ha venido en aumento la cantidad de revistas que ingresan al catálogo, así como al directorio.

⁹El Portal de revistas fue lanzado con 17 revistas que aceptaron comenzar a usar OJS. Poco más de un año después, ya se habían incluido 35 revistas restantes.

Gráfico 1



Fuente: www.latindex.org. Diciembre, 2013

Este aumento se debe básicamente a la respuesta positiva que han tenido los editores de revistas ante los talleres sobre los distintos sistemas de evaluación, sobre todo al conocer y aplicar los criterios Latindex para publicaciones impresas y electrónicas.

En este sentido, las exigencias del índice UCRIndex y las distintas capacitaciones mencionadas han logrado también una importante evolución en los criterios Latindex, por ejemplo, desde años anteriores y según evidencia las evaluaciones del 2013, cada vez más revistas cumplen con los puntos referidos a la apertura editorial, recurrir evaluadores y autores externos, y cumplimiento de la periodicidad, criterios en los que presentan las mayores tasas de incumplimiento.

Además de los criterios mencionados anteriormente, se incentiva la pertenencia a índices y bases de datos, la presencia en web, la publicación de una cantidad mínima de artículos anuales para asegurar una continuidad en el trabajo editorial de cada una de las revistas y ayuda a fomentar un grupo de lectores asiduos; y la definición de una periodicidad semestral, trimestral o menor. (UCRIndex, 2013).

Aumento de visibilidad web

Otra de los resultados obtenidos gracias a los servicios del equipo UCRIndex, por medio de sus plataformas digitales, es el aumento en visibilidad y presencia web estimado por Webometrics. En el Ranking de Portales se pasó del puesto 79 al 60 a nivel mundial, en cuanto al Ranking de Universidades, este centro de estudios universitarios ocupa actualmente el escaño 646 globalmente, el 20 en Latinoamérica y el número 1 en Costa Rica.



Estos datos confirman el empeño de llevar al espacio digital las publicaciones académicas que se producen en la UCR para asegurar un mayor nivel de impacto de las investigaciones, especialmente de las costarricenses. Gracias a la centralización de las revistas en un solo sitio es posible integrar muchos de los criterios de calidad inmediatamente a todas las revistas y brindar un buen servicio a cada una. Dentro de las necesidades cubiertas desde el portal de revista hemos logrado incorporar identificadores únicos, se realiza actualización y mantenimiento del sistema Open Journal System, creación de bloques con información general y especializada, integración de las revistas en otros sistemas y repositorios, así como conteo de estadísticas y uso de las revistas.

Reconocimiento de necesidad de profesionalización de los editores y aumento de apoyo a las revistas

A raíz del establecimiento de criterios de calidad institucionales, las autoridades formalizaron la asignación de tiempos docentes designados para labores editoriales, para convertirlos en plazas de editores asignados según resultados del índice UCRIndex. Estas acciones, junto con la capacitación, buscan la profesionalización de la labor editorial en las revistas de la UCR, punto medular para el incremento de la calidad en las publicaciones y en el crecimiento de su presencia en índices y bases de datos.

El modelo de servicio planteado parte de que “el financiamiento es un factor clave para el éxito a largo plazo de una revista académica” (Rodríguez y Tejada, 2013) pero se debe financiar y dar un seguimiento y apoyo a los editores, permitiendo aprovechar mejor el recurso invertido y garantizando la inversión en el futuro. El otorgamiento de plazas profesionales, la capacitación y la aplicación del índice UCRIndex han creado un ambiente de crecimiento y mejora continua en las revistas académicas de la Universidad.

Optimización de la inversión de recursos destinados a revistas

La optimización en la inversión de recursos se da a nivel técnico y a nivel de capacitación de personal de revistas. A nivel técnico, se trabaja con todas las revistas en el mismo servidor, lo que significa que el mantenimiento del sistema para las más de 50 revistas se hace una sola vez. También se ahorra en personal técnico encargado de esta tarea y se puede invertir en desarrollo para la herramienta OJS. en 2011, antes de la implementación del OJS, unas 26¹⁰ revistas tenían versión electrónica con sus tareas de mantenimiento asignadas a una persona diferente. Con frecuencia, esta persona asumía el mantenimiento de la revista como un recargo y no podía actualizar la información de la revista oportunamente. Muchas de estas revistas no contaban con metadatos ni estructuras adecuadas para su adecuada localización en la web.

¹⁰ 16 revistas con versión impresa y electrónica, y 11 revistas electrónicas puras. (Equipo Latindex Costa Rica, 2011)



A nivel de capacitación, al utilizar el mismo sistema de gestión de revista, es posible dar capacitación para todos los editores simultáneamente, además ganaron autonomía para la administración de los contenidos de su revista, lo cual era imposible en los anteriores sitios ya que no estaban basados en un CMS.

Con el Portal de Revistas es posible calcular la inversión en mantenimiento web de revistas y optimizar el tiempo invertido en cada una. Actualmente, 2 tiempos profesionales y 20 horas asistente son utilizados para atender las 52 revistas. Este tiempo incluye capacitación, atención individual, subida de artículos y mantenimiento del OJS¹¹. Si lo comparamos con un servicio privado como el ofrecido por Elsevier, que básicamente es un CMS para revista y la inclusión en una de sus bases de datos, la diferencia es bastante significativa. En un reciente contacto con la empresa Elsevier, cotizó por el hospedaje de una revista 24400 EUR anuales, mientras que el servicio completo le cuesta a la UCR unos 50610 USD¹² anuales para las 52 revistas, unos 973 USD anuales por revista. Además, se comparte conocimiento y desarrollos con otras universidades públicas, lo que contribuye a una disminución mayor de los costos.

CONCLUSIONES

El modelo de servicios para la gestión de publicaciones de acceso abierto ha demostrado tener múltiples ventajas, algunas ya comentadas anteriormente. Sus principales características son su agilidad para identificar necesidades de las revistas y su flexibilidad para implementar soluciones. Al designar un equipo de trabajo para centralizar el mantenimiento de la plataforma web y resolver necesidades de capacitación, es posible quitarle a los editores parte responsabilidad del mantenimiento de las versiones electrónicas y disminuir el tiempo de las curvas de aprendizaje, lo que reduce costos en hospedaje y mantenimiento de servidores, las malas prácticas relacionados con falta de capacitación, y evita la duplicación de inversión. Asimismo, se centralizó la administración de los sitios de revistas, lo que eliminó el problema del mal mantenimiento de sitios individuales y malas prácticas web de las revistas. El modelo también permite compartir la implementación de mejoras de software para todas las publicaciones por igual, incluir cambios para cumplimiento automático de criterios de calidad, y compartir experiencias entre las revistas.

¹¹ Anteriormente, era común que los encargados informáticos de las facultades asumieran, entre todas sus tareas, actualizar la versión en línea de la revista. Así, el contenido era atendido sin ninguna prioridad y quedaba a cargo de personal sin no especializado, lo que podría significar retrasos en la publicación y una falta de descripción web del contenido.

¹² Cálculo estimado según dedicación actual de 2 tiempos profesionales y 20 horas asistente mensuales. Incluye cargas sociales y otros rubros pagados por la UCR.



La utilización de Software Libre de código abierto es fundamental para este modelo, ya que permite modificar y compartir las actualizaciones libremente, haciendo posible la colaboración entre equipos de trabajo de diferentes instituciones y a las revistas compartir entre sí. Desde la administración del portal se pueden realizar cambios para que las revistas cumplan los criterios de calidad de diferentes índices.

Este mismo modelo está en proceso de aplicarse a los repositorios institucionales con miras a probar sus viabilidad en este campo y contribuir a evitar problemas posteriores de normalización y calidad de los datos recopilados desde los portales de revistas.

A nivel político, el apoyo institucional es fundamental para iniciar un proyecto como este. Todos los servicios prestados hasta el momento son producto de la negociación con las autoridades. Si bien hay servicios que cuyos resultados son tangibles de inmediato, hay otros que se visualizan con el tiempo, y es necesario saber promoverlos antes las autoridades correspondientes. En este sentido, la promoción del acceso abierto es uno de esos valores agregados para nuestra investigación cuyos resultados se apreciarán mejor en algunos años. Esta propuesta para la distribución del conocimiento sumada este modelo de servicios son una buena combinación para a) la promoción de la cultura científica y la mejora a nivel académico y científico, ya que las revistas y sus autores se ven beneficiados de las capacitaciones en temas como gestión editorial, formas correctas de para referenciar citas bibliográficas, redacción de artículos científicos, talleres de nuevos editores, actualización profesional y tecnológica; b) mejorar la visibilidad web de la institución y la investigación regional; y c) mejorar asignación de recursos económicos por parte de la administración basada en evaluación de criterios de calidad.

REFERENCIAS

Córdoba, S. (2006). Latindex: un aporte bibliotecario para el enriquecimiento de la calidad de las revistas científicas costarricenses. IV Congreso Internacional de Bibliotecología, San José, Costa Rica, noviembre 2006. <http://hdl.handle.net/10669/648>

Equipo Latindex Costa Rica (2011). Resultados de la encuesta de las revistas de la Universidad de Costa Rica. Recuperado el 15 de julio de 2014, de http://www.latindex.ucr.ac.cr/docs/Resultados_revistas.pdf

Rodriguez Yunta, L. & Tejada Artigas, C. M. (2013). El editor técnico: Un perfil necesario para la profesionalización de la edición de revistas científicas en el entorno digital. *Anales de Documentación*, 16 (2). doi:10.6018/176391

UCRIndex (2013). *Sistema de Evaluación*. Recuperado el 15 de julio de 2014, de <http://ucrindex.ucr.ac.cr/sistema-de-evaluacion-de-revistas/>



DIFUSIÓN DE LOS RESULTADOS DE INVESTIGACIÓN EN ACCESO ABIERTO DE LA “*BIBLIOTECA DIGITAL DE ESTUDIOS FEMINISTAS Y DE GÉNERO*”

Proyecto temático del sistema Nacional de bibliotecas de la Universidad
Nacional de Colombia

*Dora Isabel Diaz Susa*¹, *Sonia Maria Valencia Grajales*²

¹Directora de la Escuela de Estudios de Género. Facultad de Ciencias Humanas. Universidad
Nacional de Colombia, doraisabeldiazsusa@yahoo.es.

²Directora Nacional de Bibliotecas. Universidad Nacional de Colombia, smavalenciag@unal.edu.co

RESUMEN

El documento presenta la experiencia colaborativa de materialización de la “***Biblioteca Digital de Estudios Feministas y de Género, Ofelia Uribe de Acosta***”, ***BDEFG***, realizada por la Escuela de Estudios de Género y del Sistema Nacional de Bibliotecas (SINAB) de la Universidad Nacional de Colombia, cuya implementación se inició en el 2011; este incluye las acciones realizadas en la consolidación de su acervo bibliográfico; los procesos y procedimientos bibliotecológicos y tecnológicos implementados por el SINAB; el trabajo colaborativo y de acompañamiento brindado por un equipo interdisciplinario de las instancias involucradas; los aciertos y enseñanzas que deja el desarrollo del proyecto, el cual se configura como un producto pionero en el país por sus características al servicio de la comunidad mundial.

En la implementación de la ***BDEFG*** convergen varios propósitos políticos: posicionar y ampliar la difusión de los estudios feministas y de género, democratizar y universalizar el acceso a la información académica y científica especializada sobre la temática como estrategia clave para la transformación de la inequidad social por razones de género, clase, etnicidad o raza en la educación superior en Colombia y en la sociedad en general, gracias al aprovechamiento de las tecnologías de información (Tics) que se viene haciendo por parte de las bibliotecas del mundo, en particular de los repositorios universitarios. La ***BDEFG*** hace parte del proyecto macro “**Fortalecimiento de Capacidades relacionadas con la Equidad de Género en la Educación Superior – FEGES**”, que venimos desarrollando cuatro instituciones colombianas de educación superior: la Universidad Industrial de Santander, la Fundación Universidad Autónoma de Colombia, y la Fundación Universidad Central con la subvención del Gobierno de los Países Bajos, 2011 – 2014.

Palabras Clave: Estudios feministas y de género, acceso libre, acción colaborativa



ABSTRACT

This paper presents the collaborative experience towards the realization of the "Digital Library of Feminist and Gender Studies, Ofelia Uribe de Acosta" BDEFG, conducted by the School of Gender Studies and the National Library System (SINAB, for its Spanish acronym) of the Universidad Nacional de Colombia, whose implementation started in 2011; this includes actions to consolidate its bibliographic collection; the technological processes and library-related procedures implemented by the SINAB; collaborative and accompaniment provided by an interdisciplinary team involved in the consulting work; successes and lessons from the project, which is considered as a pioneering product in the country for its features to serve the world community. In the implementation of the BDEFG converge several political purposes: to position and to expand the dissemination of feminist and gender studies, democratize and universalize the access to academic and scientific specialized information on the topic, as a strategy for the transformation of social inequity gender, class, ethnicity or race in higher education in Colombia and in the society in general, by leveraging information technologies (ICTs) which has been done by libraries around the world, in particular with the university academic repositories. The BDEFG project is part of the macro-project "Capacity-building related to Gender Equity in Higher Education - FEGES", that we have developed with four Colombian higher education institutions: Universidad Industrial de Santander, Fundación Universidad Autónoma de Colombia, and Fundación Universidad Central with funding support by the Government of the Netherlands, 2011-2014.

Key words: Feminist and gender studies, open access, collaborative action

INTRODUCCIÓN

La propuesta de **Biblioteca Digital de Estudios Feministas y de Género (en adelante BDEFG)** hace parte del proyecto "Fortalecimiento de Capacidades relacionadas con la Equidad de Género en la Educación Superior en Colombia – FEGES", denominado NICHE/COL/061/091; el cual cuenta con la subvención de la Organización Neerlandesa para la Cooperación Internacional en Educación Superior, NUFFIC, según comunicación C&B/CPrk/6351-10 del 20 de diciembre de 2010.

Para el desarrollo del proyecto macro, en el que se incluye la presente propuesta, fueron suscritos los siguientes convenios nacionales e internacionales, que garantizan compromisos institucionales e interinstitucionales para su creación y parte de su sostenibilidad. Como primera medida debe mencionarse el convenio nacional de cooperación interinstitucional, denominado *Alianza Colombiana*, suscrito por la Universidad Nacional de Colombia, la Universidad Industrial de Santander, la Fundación Universidad Autónoma de Colombia, y la Fundación Universidad Central; esta última como representante de la *Alianza* y concesionaria del subsidio otorgado por NUFFIC. En el ámbito internacional el convenio para la asesoría general del proyecto del Consorcio Holandés entre la Universidad de Utrecht y MDF Training Consultancy, este último su representante.



La *BDEFG* se constituirá, de manera primordial, a partir de un estudio y selección del acervo bibliográfico (más de 9.000 registros especializados) del Fondo de documentación “Mujer y Género Ofelia Uribe de Acosta” de la Escuela de Estudios de Género de la Universidad Nacional de Colombia, sede Bogotá. En este proceso se incluirá también la bibliografía relevante que exista sobre el tema en las bibliotecas o fondos de documentación de las otras tres universidades de la *Alianza Colombiana*.

JUSTIFICACIÓN

En concordancia con las nuevas perspectivas de la era de la información, las bibliotecas del mundo entero están dando el paso a la digitalización de sus documentos además de diligenciar la inclusión, en sus repositorios, de documentos que respondiendo a la época, solo se producen de manera digital¹. Este cambio paradigmático en el campo de la información es sin lugar a dudas un derrotero al que se enfrentan todos los centros de documentación que, como el Fondo de Documentación Mujer y Género “Ofelia Uribe de Acosta” (FDMG), se trazaron el objetivo de constituir acervos documentales que democratizaran el acceso a la producción de conocimientos.

Entorno a temáticas que representan la búsqueda transformación de realidades inequitativas, bien sea desde el quehacer académico, o desde las luchas políticas de diferentes colectivos sociales.

Es por ello que el proyecto de *Biblioteca Digital de Estudios Feministas y de Género (BDEFG)*, constituye la oportunidad para hacer uso de las Tecnologías de la Información (TIC's) con el fin de posicionar y ampliar la difusión y al universalizar el acceso a la *información académica y científica* especializada sobre la temática, contribuir con su democratización como estrategia clave para la transformación de la inequidad social por razones de género, clase, etnicidad o raza en la sociedad en general, y de manera especial, en la educación superior en Colombia.

OBJETIVOS DEL PROYECTO

Objetivo general

Ampliar la difusión y el acceso a la producción bibliográfica relevante de los Estudios de mujeres, de género y feministas, por medio de la creación de la Biblioteca Digital de Estudios Feministas y de Género (BDEFG) a partir del acervo bibliográfico del **Fondo de Documentación Mujer y Género “Ofelia Uribe de Acosta”**, de la Escuela de Estudios de Género de la Universidad Nacional de Colombia, y de bibliografía pertinente de las bibliotecas y/o fondos de documentación de la Fundación Universidad Central, Universidad Industrial de Santander y Fundación Universidad Autónoma de Colombia, **como aporte al fortalecimiento de la equidad de género en las actividades misionales formativas, investigativas y de extensión de las universidades mencionadas.**



Objetivos específicos

- a) Seleccionar material bibliográfico con relevancia histórica, asociado a investigaciones y medios de divulgación de movimientos feministas y de mujeres de las primeras décadas del siglo XX en Colombia y América Latina y del Caribe, disponible en el Fondo de Documentación de Mujer y Género “Ofelia Uribe de Acosta” de la UN y en las bibliotecas y centros de pensamiento de las demás universidades de la Alianza colombiana participantes del convenio
- b) Seleccionar el material bibliográfico (libros, artículos de revistas y compilaciones, publicaciones seriadas, colecciones completas de revistas) que constituya un referente teórico, investigativo, o informativo destacado sobre el tema en Colombia y América Latina, en la segunda mitad del siglo XX y el presente siglo, con énfasis en la producción nacional y regional.
- c) -Digitalizar y publicar el material seleccionado en la Biblioteca Digital UN, de acuerdo con la reglamentación existente en términos de derechos de autor, y la normatividad que rige este espacio digital institucional.
- d) Creación de un dominio propio en la WEB institucional de la Universidad Nacional de Colombia, que facilite la búsqueda y acceso de las y los usuarios al material disponible en la Biblioteca Digital UN de esta biblioteca temática, con un enlace para su consulta en cada una de las universidades participantes del convenio.

METODOLOGÍA

La metodología comprende reflexiones sobre las fases del proceso de construcción de la BDEFG; la experiencia colaborativa en dialogo de saberes interdisciplinarios de Estudios feministas y de género, la bibliotecología y las Tics, con diferentes niveles de coordinación y cooperación intrainstitucional, interinstitucional, nacional e internacional; los criterios y la forma de selección del acervo bibliográfico por su relevancia histórica, teórica, investigativa, o informativa, asociada a investigaciones y medios de divulgación de movimientos feministas y de mujeres desde las primeras décadas del siglo XX hasta la actualidad en Colombia, América Latina y el Caribe; Igualmente, comprende la apropiación y adaptación de los procedimientos elaborados por el Sinab para el montaje del repositorio institucional en la biblioteca digital y las bibliotecas temáticas en lo atinente a digitalización, publicación de acuerdo con la reglamentación existente en términos de derechos de autor, y a la normatividad que rige este espacio digital institucional e internacional. Para cada uno de esos aspectos se señalan posibilidades y limitaciones.



RESULTADOS PARCIALES

Tomando como base el Fondo de documentación “Mujer y Género Ofelia Uribe de Acosta”, pionero en la temática en el país (1986) y la bibliografía relevante que existe sobre el tema en las bibliotecas o fondos de documentación de las otras tres universidades participantes en el proyecto *FEGES*; este acervo bibliográfico se amplía con nuevas autorizaciones de libre acceso, logradas en las gestiones con diferentes autoras y autores.

Repositorio digital con metadatos estandarizados bajo las directrices DRIVER – OPEN AIRE, con lenguaje especializado en el área, con herramienta de cosechador para integrarse a otros proyectos mundiales de similares características.

Acceso a un acervo bibliográfico de al menos 1000 títulos especializados.

Página web con funcionalidades adecuadas para una búsqueda eficiente y eficaz.

Reflexiones finales

Este trabajo interdisciplinario e interinstitucional es un ejemplo de las posibilidades que existen de visibilizar la ciencia y la investigación de un país en un área específica y de tanta importancia para el mundo entero.

BIBLIOGRAFÍA

Biblioteca Nacional de Colombia. Directrices de digitalización.

Disponible en

<http://www.bibliotecanacional.gov.co/index.php?idcategoria=39244>

Biblioteca Nacional de Colombia. Pautas para la digitalización de documentos análogos en papel y audio. En: *Conservamos*, no. 7, ene-feb, 2009. Bogotá, Colombia.

Consorti de Biblioteques Universiàries de Catalunya. Estàndars de digitalització. Elementos mínimos. Disponible en <<http://www.cult.gva.es/DAIT/Archivos/images-reino/Recdigitalizacion.pdf>>

Dirección Nacional de Bibliotecas. Perfiles del equipo de trabajo de la Biblioteca Digital UN en el Sistema Nacional de Bibliotecas (SINAB) versión 1.0

Dirección Nacional de Bibliotecas. Guía para la formulación y desarrollo de proyectos temáticos de Biblioteca Digital UN (SINAB) versión 1.0



IFLA e ICA. Directrices para proyectos de digitalización. Disponible en
<http://travesia.mcu.es/documentos/pautas_digitalizacion.pdf>

Library of Congress. Digital Formats for Content Reproductions: Disponible en
<<http://memory.loc.gov/ammem/formats.html>>

Library of Congress: Manuscript Digitization Demonstration Project. Disponible en:
<<http://memory.loc.gov/ammem/pictel/>>

Universidad de Cornell. Llevando la teoría a la práctica. Tutorial de Digitalización de imágenes. Disponible en: < <http://www.library.cornell.edu/preservation/tutorial-spanish/contents.html>>



Sessão de pôsters

Apresentados na sessão “Meu pôster em um minuto”



O REPOSITÓRIO ISCTE-IUL E A INTEROPERABILIDADE COM OUTROS SISTEMAS DE INFORMAÇÃO

*Maria João Amante, Teresa Maria Segurado, Bruno Marçal
ISCTE-Instituto Universitário de Lisboa*

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/98/124

Resumo

ISCTE-IUL, a través de sus Servicios de Información y Documentación ha realizado una apuesta muy clara en la promoción del Acceso Libre al Conocimiento traducida en distintas acciones realizadas y también a través de la adopción de las herramientas tecnológicas las más adecuadas y de las políticas institucionales necesarias para desarrollar distintos proyectos en este dominio. En este contexto, no podemos dejar de destacar que hemos sido una de las primeras instituciones de enseñanza superior en Portugal a desarrollar un repositorio que destinado a gestionar la producción científica desarrollada por sus profesores e investigadores (en el año del 2006), contribuyendo no apenas para su preservación pero también para su divulgación y proyección, nacional e internacional, a la cual muy ha contribuido a su política de auto-archivo, adoptada en noviembre del 2009. Este esfuerzo es desarrollado, de forma continua, en distintos locales y momentos, por ejemplo, a través de la promoción de la plataforma junto de los usuarios internos y externos, de actividades de atracción de la comunidad académica como estrategia de motivación al depósito de documentos, de la participación en proyectos de ámbito nacional (como el Portal RCAAP o en actividades más globales como la Semana Internacional del Acceso Libre. Existe igualmente una preocupación muy clara con la garantía de que el Repositorio mantiene la conformidad con las directivas existentes al nivel internacional para este tipo de plataformas. El Repositorio ISCTE-IUL se ha establecido como una fuente más de información a la disposición de toda la comunidad científica, marcado por un crecimiento sólido de los indicadores cuantitativos a lo largo de los años reforzado por la manifiesta confianza en la calidad de sus contenidos por parte de todos aquellos que lo consultan. ISCTE-IUL ha recientemente adquirido un Portal Agregador de Contenidos (Retrievo) con el objetivo ayudar sus usuarios a encontrar información relevante, disponible en distintas fuentes, de forma sencilla, rápida y eficiente. Una de las condiciones para la selección de la solución era la garantía de precisamente pela garantía da interoperabilidade de la herramienta a adquirir con el Repositorio a través del protocolo OAI-PMH. El repositorio de la institución tiene, en este momento, 6200 documentos depositados, la mayoría en acceso libre (76%), distribuidos por 34 comunidades y 384 colecciones (que reflejen la a estructura orgánica actual de la institución). El poster nos proponemos presentar tiene como objetivos principales objetivos dar a conocer el Repositorio ISCTE-IUL, su tasa de crecimiento/utilización, y mostrar de qué forma los usuarios acceden a la información haciendo uso de distintas plataformas, contribuyendo para aumentar la calidad de la actividad científica y académica al mismo tiempo que se simplifican procesos y se disminuye la dispersión de la información.

Palavras-chave: Acceso libre; Repositorio ISCTE-IUL; Portal Agregador de Contenidos; Interoperabilidad; OAI-PMH.



USABILIDADE EM PERIÓDICOS ELETRÔNICOS DE ACESSO ABERTO: O CASO DA REVISTA EM QUESTÃO

Gonzalo Rubén Alvarez¹, Juliana Guerra de Barcellos², Samile Andréa de Souza Vanz³; Helen Beatriz Frota Rozados⁴

¹Mestrando do PPGCOM/FABICO/UFRGS; ²Graduanda de Biblioteconomia da FABICO/UFRGS; ³Professora do Programa de Pós-Graduação em Comunicação e Informação da FABICO/UFRGS; ⁴Professora Associada do Departamento de Ciências da Informação da FABICO/UFRGS

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/105/108

Resumo

A inclusão de sofisticadas ferramentas de comunicação da informação no meio digital causou uma demanda por produtos altamente adaptáveis às necessidades dos seus usuários. Nesse sentido, a usabilidade tornou-se essencial porque permite avaliar o grau de qualidade desses produtos de acordo com a sua facilidade de uso. Os *designers* de produtos buscam aproximar a tecnologia com os usuários dos mesmos e adaptá-los de acordo com os objetivos para os quais são usados. Como canais formais de informação, os periódicos eletrônicos de acesso livre são muito utilizados porque oferecem rapidez na comunicação das pesquisas e dissolvem as barreiras de acessibilidade. *Open Access* estimula a participação da comunidade científica para divulgar suas investigações em periódicos e repositórios institucionais de acesso livre, com a finalidade de alcançar maior visibilidade, alargar o diálogo e a parceria entre seus membros. A publicação das pesquisas é vital para o crescimento da ciência e para o desenvolvimento das diferentes áreas. Entretanto, para alcançar um nível elevado de eficiência na publicação, os *softwares* que gerenciam o processo editorial devem apresentar interfaces simples e intuitivas. Interfaces com essas características produzem satisfação e motivam os usuários para continuarem utilizando o produto. Diversos projetos foram desenvolvidos para incentivar a editoração eletrônica de revistas científicas. No Brasil, o Sistema Eletrônico de Editoração de Revistas (SEER), traduzido do *Open Journal System* (OJS) pelo IBICT, é um dos *softwares* mais utilizados na gestão de periódicos eletrônicos, sendo o encarregado da automatização das atividades de editoração. A revista *Em Questão* é uma das publicações da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul que utiliza o SEER e foi objeto de análise de uma pesquisa acerca da usabilidade da seção de submissão de artigos. Participaram da pesquisa 69 autores com o objetivo de avaliar a usabilidade, com base nos cinco atributos propostos por Jakob Nielsen: Facilidade de manuseio; Eficiência de uso; Facilidade de memorização; Baixa taxa de erros; Satisfação do usuário. A partir das diversas recomendações da pesquisa foram propostas algumas alterações na revista *Em Questão*, como as modificações em alguns atributos: Facilidade de manuseio (Realces: Cores), (Legibilidade: Tamanho da fonte, Espaçamento entre linhas, Espaçamento entre parágrafos); Facilidade de memorização (Dificuldade para memorizar as etapas do processo de submissão); Taxa de erros (Apresentação de mensagens de erros diante uma ação); a



personalização da senha de acesso ao sistema; migração para uma versão mais atualizada do software SEER; personalização das páginas e dos botões; utilização de *templates* atrativos; destaque do ISSN, *layout* de três colunas, destaque da equipe editorial e modificações CSS. Este trabalho apresenta o relato das modificações realizadas na Em Questão no que diz respeito à usabilidade da revista e propõe alguns padrões de qualidade que podem ser aplicados em revistas que utilizam o SEER.

Palavras-chave: Open Access; Periódico científico eletrônico; Sistema Eletrônico de Editoração de Revistas (SEER); Usabilidade; Revista Em Questão.

Referências

ALVAREZ, G. R. **Usabilidade da seção de submissão do Sistema Eletrônico de Editoração de Revistas:** uma análise a partir da opinião dos autores cadastrados na revista Em Questão. 2013. 112 f. Trabalho de Conclusão de Curso (Graduação)– Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013.

EM QUESTÃO: Revista da Faculdade de Biblioteconomia e Comunicação. Porto Alegre: Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, 2014 – Semestral.

GRUSZYNSKI, A. C.; CASTEDO, R. O *design* de periódicos científicos *on-line* e a visibilidade da ciência na web. **InfoDesign**, São Paulo, v. 6, n. 3, p. 1-11, 2009.

USABILIDAD EN PERIÓDICOS ELECTRÓNICOS DE ACCESO ABIERTO: EL CASO DE LA REVISTA EM QUESTÃO

Gonzalo Rubén Alvarez, Juliana Guerra de Barcellos, Helen Beatriz Frota Rozados, Samile Andréa de Souza Vanz

Resumen

La inclusión de sofisticadas herramientas de comunicación de la información en el medio digital ha provocado una demanda por productos altamente adaptables a las necesidades de sus usuarios. De acuerdo con ello, la usabilidad se volvió esencial ya que permite evaluar el grado de calidad de esos productos en función de su facilidad de uso. Los diseñadores de productos buscan aproximar la tecnología con los usuarios de los mismos y adaptarlos de acuerdo con los objetivos para los que se utilizan. Como canales formales de información, los periódicos electrónicos de acceso abierto son muy utilizados porque ofrecen rapidez en la comunicación de los estudios y disuelven las barreras de accesibilidad. *Open Access* fomenta la



participación de la comunidad científica para divulgar sus investigaciones en periódicos y repositorios institucionales de acceso abierto, con el fin de lograr una mayor visibilidad, ampliar el diálogo y la colaboración entre sus miembros. La publicación de la investigación es vital para el crecimiento de la ciencia y para el desarrollo de las diferentes áreas. Sin embargo, para alcanzar un nivel elevado de eficiencia en la publicación, los *softwares* que gerencian el proceso editorial deben proporcionar interfaces simples e intuitivas. Interfaces con esas características producen satisfacción y motivan a los usuarios para continuar utilizando el producto. Se han desarrollado varios proyectos para incentivar la publicación electrónica de revistas científicas. En Brasil, el Sistema Electrónico de Edición de Revistas (SEER), traducido del *Open Journal System* (OJS) por el IBICT, es uno de los *softwares* más utilizados en la gestión de periódicos electrónicos, estando a cargo de la automatización de las actividades de edición. La revista *Em Questão* es una de las publicaciones de la *Faculdade de Biblioteconomia e Comunicação* de la *Universidade Federal do Rio Grande do Sul* que utiliza el SEER y fue objeto de análisis de un estudio acerca de la usabilidad de la sección de sumisión de artículos. Participaron del estudio 69 autores con el objetivo de evaluar la usabilidad, con base en los cinco atributos propuestos por Jakob Nielsen: Facilidad de manejo; Eficiencia de uso; Facilidad de memorización; Baja tasa de errores; Satisfacción del usuario. A partir de las diversas recomendaciones del estudio, fueron propuestas algunas alteraciones en la revista *Em Questão*, como las modificaciones en algunos atributos: Facilidad de manejo (Realces: Colores), (Legibilidad: Tamaño de la letra, Espaciamiento entre las líneas, Espaciamento entre párrafos); Facilidad de memorización (Dificultad para memorizar las etapas del proceso de sumisión); Tasa de errores (Presentación de mensajes de error delante de una acción); la personalización de la clave para acceder al sistema; la migración para una versión más actualizada del *software* SEER; personalización de las páginas y de los botones, utilización de *templates* atractivos; destaque del ISSN, *layout* de tres columnas, destaque del equipo editorial y modificaciones CSS. Este trabajo presenta una descripción de los cambios realizados en la *Em Questão* con respecto a la usabilidad de la revista y propone algunos padrones de calidad que pueden ser aplicados en revistas que utilizan el SEER.

Palabras clave: *Open Access*. Periódico científico electrónico. Sistema Electrónico de Edición de Revistas (SEER). Usabilidad. Revista *Em Questão*.



MEJORAMIENTO DE LA VISIBILIDAD DE LA PRODUCCIÓN CIENTÍFICA EN EL INSTITUTO TECNOLÓGICO DE COSTA RICA

Alexa Ramírez-Vega
Instituto Tecnológico de Costa Rica

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/106/129

Resumen

La utilización de Internet, y especialmente de la Web 2.0, como medio principal de intercambio de información ha hecho que cualquier persona pueda tener acceso a la producción científica, proveniente de instituciones educativas, centro de investigación y academias en todo el mundo. Esta producción es uno de los indicadores clave en el momento de evaluar el grado de excelencia y desarrollo de la generación de conocimiento de una institución o país determinado (Machín, Lugo, & Díaz, 2013). Por lo tanto, la divulgación de los resultados de investigación adquiere mayor relevancia para los investigadores, quienes además de concluir satisfactoriamente los proyectos de investigación, deben dar a conocer los aspectos más relevantes de su labor y así contribuir al mejoramiento y aumento de la producción científica de su institución. Por su parte, para medir la producción científica se cuenta con variados indicadores, que van desde la incorporación en índices y bases de datos de revistas, rankings mundiales de universidades, hasta la presencia en la web. En todos ellos la visibilidad juega un papel predominante, de manera que la inclusión de las publicaciones periódicas en índices y bases de datos reconocidas, además de prestigio, brinda mayor visibilidad a los documentos que ahí se publican; de igual forma los rankings de universidades como el *Webometrics Rankings of World Universities* del Laboratorio de Cibermetría del Consejo Superior de Investigación Científica de España (CSIC, España, 2014) es calculado basado principalmente en el contenido presente en la web, proveniente de diferentes bases de datos académicas; finalmente, la presencia en la web consistirá de la cantidad de contenidos disponibles en Internet y la posibilidad de ser recuperados por buscadores especializados. De esta manera, la visibilidad de los contenidos en la web deben poseer dos características: disponibilidad y confiabilidad. La disponibilidad en cuanto el contenido pueda ser recuperado por cualquier lector en la web; y confiabilidad en cuanto el contenido recuperado sea catalogado como académico y se logre identificar su autor y procedencia. Dado lo anterior, el Instituto Tecnológico de Costa Rica (ITCR) inició varios procesos para aumentar la visibilidad de sus resultados de investigación a través de la vía verde y la vía dorada del Acceso Abierto (Hernández Pérez, Rodríguez Mateos & Bueno De la Fuente, 2007), con el fin de aumentar la cantidad de contenidos en la web, garantizar su disponibilidad y mejor la confiabilidad de éstos. La principal problemática encontrada fue la escasa presencia de contenidos en la web, lo cual se trató de remediar con la creación del portal de revistas académicas del TEC (Ramírez-Vega, 2013). También, se inició el proceso para fortalecer el repositorio institucional a través de su mejora y la creación de políticas institucionales que agilicen la incorporación de contenidos. Además, se realizaron talleres a investigadores del ITCR, donde se les capacitó en



el uso de herramientas (Google Scholar y redes sociales académicas) para mejorar su visibilidad. Así mismo, se impartieron charlas que permitieron identificar problemáticas e inconsistencias en los nombres de las publicaciones y la poca normalización en el uso del nombre de los autores y el de la institución, encontrando hasta cinco formas distintas para el ITCR. En este trabajo se exponen los resultados obtenidos con la aplicación de las estrategias antes mencionadas, de forma que se evidencien las debilidades encontradas, las posibilidades de mejora y las acciones correctivas que ya fueron puestas en práctica, las cuales han permitido el mejoramiento de la visibilidad de la producción de los investigadores, de la institución y del país en general.

Referencias

CSIC, España. (2014). Webometrics Rankings of World Universities. Retrieved May 16, 2014, from <http://www.webometrics.info>

Hernández Pérez, T., Rodríguez Mateos, D., & Bueno De la Fuente, G. (2007). Open Access: el papel de las bibliotecas en los repositorios institucionales de acceso abierto. Retrieved from <http://digitum.um.es/xmlui/handle/10201/4017>

Machín, M. P., Lugo, Y. G., & Díaz, M. del P. F. (2013). Análisis bibliométrico sobre la visibilidad científica y el factor de impacto en ciencias de la educación. *EDUMECENTRO*, 3(3), 99–112.

Ramírez-Vega, A. (2013). Desarrollo e impacto del portal de revistas electrónicas del TEC. *Investiga.TEC*, (18), pág. 16–19.



ANÁLISE DAS CITAÇÕES DA REVISTA EM QUESTÃO ATRAVÉS DO GOOGLE ACADÊMICO

Suelen Spíndola Bilhar¹; Samile Andréa de Souza Vanz²

¹Graduada em Biblioteconomia pela Universidade Federal do Rio Grande do Sul; ²Profª. Drª. do Departamento de Ciências da Informação e do Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/112/130

Resumo

O pôster apresenta resultados da análise das citações à revista de acesso aberto Em Questão no período de 2003 a 2010. A análise de citações permite subsidiar e dar apoio à tomada de decisões visando a maior qualificação do periódico, além de possibilitar aos potenciais autores o prévio conhecimento da visibilidade que o artigo terá. Para o alcance do objetivo foram identificadas as publicações e as citações recebidas no período. A análise das citações recebidas pela revista permitiu a avaliação da visibilidade dos autores e dos artigos, através da identificação dos autores, periódicos e eventos citantes, idioma dos documentos citantes e o tempo de citação. O trabalho se fundamenta em estudos teóricos acerca da comunicação científica, periódicos científicos e suas avaliações, estudos bibliométricos e análise de citação e, por fim, o uso do *Google Acadêmico* como ferramenta para o estudo de citação. Com o intermédio do *software Publish or Perish*, o *Google Acadêmico* foi utilizado como fonte de pesquisa para a coleta dos dados feita entre abril e maio de 2013. Entre os resultados mais significativos da pesquisa, destaca-se que das 190 publicações do período analisado, o ano de 2010 foi o mais produtivo em termos de artigos, porém, o ano de 2006 foi o mais impactante, pois as publicações desse ano receberam 184 citações. As publicações aumentaram anualmente no período analisado. O artigo mais citado é uma revisão intitulada “Bibliometria: evolução histórica e questões atuais”, que recebeu 114 citações. Os três artigos mais citados pertencem à área da Ciência da Informação e são responsáveis, juntos, por 53% do total das citações. Observou-se que os autores mais produtivos não são os mais citados. Das 528 citações recebidas, 223 estavam em artigos de periódicos e 114 em trabalhos apresentados em eventos. O tempo decorrido entre a publicação e citação é predominantemente de seis anos. O trabalho concluiu que o *Google Acadêmico* serve como instrumento para análises bibliométricas, possibilitando, desta forma, a análise de citações para as revistas de acesso aberto não indexadas pela *Web of Science* ou *Scopus*. A partir dos resultados obtidos sugere-se que as revistas de acesso aberto realizem periodicamente análise de citação para acompanhamento de sua visibilidade entre a comunidade científica. A pesquisa também indicou que novas estratégias de *marketing* podem ser elaboradas para a revista Em Questão, bem como salientou que esforços devem ser somados para



que parâmetros sejam seguidos com vistas à melhor difusão entre a comunidade científica e a melhor avaliação do periódico pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Palavras-chave: Bibliometria; Análise de citação; Visibilidade; Periódico científico; Em Questão; Google Acadêmico; Publish or Perish

Referências

CAREGNATO, Sônia Elisa. Google acadêmico como ferramenta para estudos de citações: avaliação da precisão das buscas por autor. **PontodeAcesso**, Salvador, v. 5, n. 3, dez. 2011.

MACIAS-CHAPULA, Cesar A. O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. **Ci. Inf.**, Brasília, DF, v. 27, n. 2, maio/ago. 1998.

MEADOWS, Arthur Jack. **A comunicação científica**. Brasília, DF: Briquet de Lemos, 1999.

MUGNAINI, Rogério; STREHL, Leticia. Recuperação e impacto da produção científica na Era Google: uma análise comparativa entre o Google Acadêmico e a Web of Science. **Encontros Bibli**, Florianópolis, n. esp., 2008.

SOLLA PRICE, Derek J. Society's need in scientific and technical information. **Ci. Inf.**, Brasília, v. 3, n. 2, p. 97-103, 1974.



INCLUSÃO DE CONTEXTO EM ARTIGOS CIENTÍFICOS: EM BUSCA DA WEB PRAGMÁTICA

Denysson Mota¹, Nair Yumiko Kobash²

¹Doutorando em Ciência da Informação – Universidade de São Paulo; ²Doutora em Ciências da Comunicação pela Universidade de São Paulo (1994). Professora do Departamento de Biblioteconomia e Documentação e do Programa de Pós-graduação em Ciência da Informação da Universidade de São Paulo

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/133/115

Resumo

Este trabalho discute a Interoperabilidade e a Recuperação de Informação (RI) em bibliotecas digitais, questões que desafiam a Web atual. A proposta da Web Semântica, de descrever diferentes recursos por meio de metadados tem promovido avanços na RI. Porém é possível ir além. O passo seguinte seria adicionar contexto aos recursos, para caminhar em direção a uma Web Pragmática, com base nas teorias semiótico-pragmáticas de Peirce (1977) e Morris (1938). Esta perspectiva vem sendo discutida por autores como Signh (2002) e de Moor, Keeler e Richmond (2002), assim como por Sowa (2000) e Ingwersen e Järvelin (2005), que apontam os problemas teóricos e tecnológicos das ferramentas semânticas atuais. Os conceitos da Web Pragmática vêm sendo aplicados em ambientes comerciais, mas não são objeto de pesquisas de aplicação em bibliotecas digitais acadêmicas. Neste texto serão apresentadas as propostas da Web Pragmática no que concerne as tecnologias de integração de contexto ao conteúdo de trabalhos científicos, como o RDF (W3C, 2006) e os Topic Maps (ISO, 2003). Os Topic Maps, forma de representação da informação (ISO/IEC 13250, 2003), têm a capacidade de definir conceitos e suas relações dentro de escopos específicos (ISO, 2006), ampliando as possibilidades de inclusão de contexto aos dados. Podem ser, assim, criadas bases bibliográficas semântica e pragmaticamente mais ricas que aprimoram os resultados da RI. Na presente etapa da pesquisa, está em análise um repositório brasileiro de artigos de periódicos científicos da área da Ciência da Informação, cujos registros estão sendo comparados com registros bibliográficos adaptados para a Web Semântica, como o da British Library. Os resultados preliminares do uso dos Topic Maps, na sua forma expressa em XML (ISO, 2013), mostram a viabilidade da aplicação a grandes bases bibliográficas científicas, com resultados promissores.

Palavras-chave: RDF. Recuperação da Informação. Repositórios Digitais. Web Pragmática. Web Semântica. Topic Maps



Referências

DE MOOR, A.; KEELER, M.; RICHMOND, G. Towards a Pragmatic Web. Conceptual Structures: Integration and Interfaces. **Lecture Notes in Computer Science**, v. 2393, 2002, pp. 235-249.

INGWERSEN, P. **Information Retrieval Interaction**. London: Taylor Graham, 1992.

INGWERSEN, P.; JÄRVELIN, K. **The Turn: Integration of Information Seeking and Retrieval in Context**. The Netherlands: Springer, 2005.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). ISO/IEC 13250:2003: **Information technology -- SGML applications -- Topic maps**. Genebra, 2003.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). ISO/IEC 13250-2:2006: **Information technology -- Topic Maps -- Part 2: Data model**. Genebra, 2006.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). ISO/IEC 13250-3:2013: **Information technology -- Topic Maps -- Part 3: XML syntax**. Genebra, 2013.

PEIRCE, C. S. **Semiótica**. São Paulo: Perspectiva, 1977.

MORRIS, C. **Foundations of the Theory of Signs**. Chicago: Univ. of Chicago Press, 1938.

SINGH, M.P. The Pragmatic Web: Preliminary thoughts. In: **Proceedings of the NSF-OntoWeb Workshop on Database and Information Systems Research for Semantic Web and Enterprises**, 2002, pp. 82–90.

SOWA, F. **Knowledge Representation: Logical, Philosophical, and Computational Foundations**. Pacific Grove: Brooks/Cole., 2000.

WORLD WIDE WEB CONSORTIUM (W3C). **Defining N-ary Relations on the Semantic Web**. 2006. Disponível em < <http://www.w3.org/TR/swbp-n-aryRelations/> >. Acesso em < 20 abr. 2014 >

WORLD WIDE WEB CONSORTIUM (W3C). **Resource Description Framework (RDF)**. 2014. Disponível em < <http://www.w3.org/RDF/> >. Acesso em < 20 abr. 2014 >



INOVAÇÃO COLABORATIVA E REPOSITÓRIOS DE DADOS CIENTÍFICOS: PROPOSTA DE AMBIÊNCIA PARA PROMOÇÃO DE TECNOLOGIAS PARA A AGRICULTURA FAMILIAR

Daniela Maciel Pinto

Graduada em Biblioteconomia e Documentação pela Universidade de São Paulo

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/134/116

Resumo

Diante das questões sobre a transferência de informação tecnológica para agricultores familiares é cada vez mais frequente o registro de experiências que propõem mudança de enfoque na geração de tecnologias e que abrangem a discussão sobre novos métodos, sendo que seus resultados enriquecem o debate sobre a construção de modelos inovadores. O processo de inovação tecnológica é fundamental para o desenvolvimento da agricultura nacional e, atualmente, tem sido visto de uma maneira ampla, integrando ao processo agentes internos e externos de uma organização (BUENO; BALESTRINI, 2012). Numa sociedade em que o imperativo é conhecer e ter maior capital intelectual (TOMAEL et al., 2005), abrir-se ao conhecimento externo é essencial às instituições, principalmente quando o amadurecimento das TICs, especialmente voltadas para a colaboração em massa, permitem a interação e criação com mais indivíduos em mais regiões do mundo, usando, por isto, um conjunto de capacidades mais versátil. A construção compartilhada, pensada para o processo de inovação tecnológica na agricultura, deve ser vista como um modelo capaz de romper as fronteiras organizacionais, criando conexões entre todos os atores envolvidos no processo: consumidores, parceiros, fornecedores e demais agentes, para agregar valor ao produto final. Um projeto da Embrapa tem como objetivo estruturar uma ambiência, que propõe a produção coletiva do conhecimento (crowdsourcing) a partir de serviços de informação tecnológica como repositórios de dados científicos de acesso aberto, como um processo alternativo e auxiliar de geração de novas tecnologias para a agricultura familiar. Para subsidiar o projeto, realizou-se um levantamento na base de dados "Agrícola", sobre o tema da colaboratividade e inovação e discutiu-se dois trabalhos, sobre estruturas de Serviços de Informação Tecnológica, onde ocorre a coleta e registro de informações oriundas dos consumidores e usuários, em que a colaboratividade tem sido vista de maneira estratégica para a resolução de questões comuns, bem como proposições de novas abordagens para a P&D.

Palavras-Chave: Inovação Aberta; Interoperabilidade; Colaboratividade, Transferência de Tecnologias; Repositórios de dados



Referências

BUENO, B.; BALESTRIN, A. Inovação colaborativa: uma abordagem aberta no desenvolvimento de novos produtos. *Revista de Administração de Empresas, São Paulo*, v. 52, n. 5, 2012. Disponível em: <http://www.scielo.br/scielo.php?pid=S003475902012000500004&script=sci_abstract&tlng=p>. Acesso em 23 nov. 2013.

TOMAÉL, M. I.; ALCARÁ, A. R.; DI CHIARA, I. G. Das redes sociais à inovação. *Ciência da Informação, Brasília*, V. 34, n. 2, p. 93-104, 2005.

COLLABORATIVE INNOVATION AND SCIENTIFIC DATA REPOSITORIES: PROPOSED OF AN AMBIENCE TO PROMOTE TECHNOLOGIES FOR FAMILY FARMS

Daniela Maciel Pinto

Graduada em Biblioteconomia e Documentação pela Universidade de São Paulo

Abstract

Faced with questions about the transfer of technological information to farmers is increasingly common logging experiments that propose change of focus on generating technologies and spanning the discussion on new methods, but results enrich the debate about building innovative models. The process of technological innovation is key to the development of national agriculture and currently has been seen in an extensive way, integrating the process internal and external stakeholders of an organization, (BUENO ; BALESTRINI 2012). In a society where the imperative is to know and have greater intellectual capital (TOMAEL et al., 2005) , opening up to external knowledge is essential to the institutions, especially when the maturity of ICTs, especially geared for mass collaboration , allow interaction and creating more individuals in more parts of the world, using, for this, a set of versatile capabilities. The shared construction, designed for the process of technological innovation in agriculture, should be seen as a model capable of breaking organizational boundaries, creating connections between all actors involved in the process: customers, partners, suppliers and other agents, to add value to final product. A project of Embrapa aims to structure an ambience, which proposes the collective production of knowledge (crowdsourcing) from information technology services, as repositories of scientific data open access, as an alternative process and help to generate new technologies for family farming. To support the project, carried out a survey in the database " Agrícola ", on the subject of colaboratividade and innovation was discussed two studies on structures of Information Technology Services, where the gathering and recording information from the occurs consumers and users, where co-creation has been seen in a strategic way to resolve common issues and proposals for new approaches to R&D.



ARQUIVOS ABERTOS NA PLATAFORMA LINKED SCIENCE: A EXPERIÊNCIA DO AMBIENTE FABRICO/CIÊNCIA

Rafael Port da Rocha¹, Isadora Turcate²

¹Professor do Departamento de Ciências da Informação da Universidade Federal do Rio Grande do Sul; ²Graduanda em Biblioteconomia pela Universidade Federal do Rio Grande do Sul

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/142/125

Resumo

A web trouxe grandes benefícios à comunicação científica, possibilitando a publicação de resultados de pesquisa de acesso livre em Arquivos Abertos, e a formação de federações destes Arquivos Abertos. A Web Semântica e os Dados Ligados (Linked Data) trazem novas possibilidades para a comunicação e o intercâmbio de recursos da ciência. Enquanto que Arquivos Abertos estão focados na publicação e na interoperabilidade de publicações científicas na web, a plataforma Web Semântica/Linked Data possibilita a interoperabilidade semântica de recursos da ciência, que engloba também outros tipos recursos, como dados de pesquisa, pesquisadores, instituições, projetos, processos e eventos. A representação de recursos da ciência na Web Semântica tem sido chamada de Linked Science, que amplia as possibilidades de trocas no desenvolvimento da ciência, promove a descoberta de recursos, e aproxima pesquisas em um ambiente distribuído. Recursos da ciência já estão presentes na Web Semântica/Linked Data, como as bases de dados bibliográficas Faceted DBLP¹ e Semantic Web Dog Food². Europeana Linked Data Pilot³ traz para a Web Semântica os registros do projeto Europeana, serviço que colhe metadados de 1500 instituições culturais da Europa. VIVO⁴ e Eagle-i⁵ são ambientes que visam promover a colaboração entre pesquisadores de diversas instituições e envolvem uma diversidade de recursos da ciência, como instituições, pesquisadores e projetos. Fabrico/Ciência é um ambiente que está sendo desenvolvido para experimentar a publicação na Web Semântica de registros bibliográficos colhidos de arquivos abertos, e sua integração com outros tipos de recursos da ciência, a fim de promover a descoberta de recursos. Isso envolve questões como: representação, mapeamento e armazenamento na Web Semântica de metadados colhidos em Arquivos Abertos; identificação e representação (via ontologias) de outros tipos de recursos da ciência; e interligação destes com os recursos colhidos, enfocando a interoperabilidade semântica e o reuso de ontologias. Este artigo apresenta os resultados obtidos no que diz respeito à publicação na Web Semântica de dados colhidos em Arquivos Abertos através do ambiente Fabrico/Ciência, e estudos em andamento referentes à representação no Fabrico/Ciência de outros recursos da ciência. Descreve a arquitetura desenvolvida para Fabrico/Ciência, apresenta os resultados obtidos através da colheita de

¹ Faceted DBLP. <http://dblp.l3s.de/d2r>

² Semantic Web Dog Food. <http://data.semanticweb.org/>

³ Europeana Linked Data Pilot <http://europeana.ontotext.com>

⁴ VIVO. <http://vivoweb.org/about>

⁵ Eagle-i. <https://www.eagle-i.net/>



metadados de uma revista eletrônica, e identifica os principais tipos de recursos da ciência presentes em VIVO e Eagle-i, a fim de subsidiar o desenvolvimento de uma ontologia para o Fabrico/Ciência. Conclui que a arquitetura desenvolvida é adequada para representar na Web Semântica recursos colhidos de arquivos abertos, que VIVO e Eagle-I descrevem recursos da ciência visando à colaboração, e que as ontologias usadas nesses projetos estão sendo combinadas através de uma estrutura semântica integrada. Propõe que seja desenvolvido para o ambiente Fabrico/Ciência uma ontologia com base nos esforços de integração entre VIVO e Eagle-i, envolvendo a descrição de recursos ligados à ciência, como pessoas (e seus papéis e relacionamentos), organizações (como universidades, institutos, fundações), grupos (como equipes e comitês), eventos (processo, projetos, conferências, cursos) e objetos que contém informação (documentos, protocolos, coleções, bases de dados).

Palavras-Chave: dados ligados, arquivos abertos, linked science

Referências

Faceted DBLP. <http://dblp.l3s.de/d2r>

Semantic Web Dog Food. <http://data.semanticweb.org/>

Europeana Linked Data Pilot <http://europeana.ontotext.com>

VIVO. <http://vivoweb.org/about>

Eagle-i. <https://www.eagle-i.net/>



ESTUDO INVESTIGATIVO SOBRE METADADOS NO CONTEXTO DAS “3 MARIAS”

Ana Carla Cunha Nascimento¹, Márcio Bezerra da Silva², Mayara Campos Souza³
¹Graduanda em Biblioteconomia pela Universidade de Brasília; ²Doutorando em Ciência da
Informação pela Universidade Federal da Bahia; ³Graduanda em Biblioteconomia pela Universidade
de Brasília

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/146/126

Resumo

Estudo teórico sobre padrões de metadados enquanto instrumentos de descrição e mediação da informação no contexto das “3 Marias” apresentado por Smit (1993). Discute brevemente as “3 Marias” enquanto áreas presentes no certame da Ciência da Informação (CI) e investiga metadados em seus conceitos e modelos. Objetiva-se de forma ampla analisar os padrões de metadados adotados pelas “3 Marias” e, especificamente, elencar padrões de metadados adotados por cada área e mapear os campos de metadados comuns entre os padrões identificados. Lança mão de um percurso metodológico a partir de estudo bibliográfico e uso da técnica de mapeamento crosswalks. Verifica como resultado de pesquisa campos semelhantes entre os padrões identificados como autoria, título e data, além de constatar que os padrões de metadados perpassam pela descrição e preservação. Conclui-se que os padrões de metadados são instrumentos de descrição e mediação que podem contribuir positivamente nos processos de recuperação da informação na Web.

Palavras-Chave: Metadados, 3 Marias, Descrição e mediação da informação

Referências

ALVES, Maria das Dores R.; SOUZA, Márcia Izabel F. Estudo de correspondência de elementos de metadados: Dublin Core e MARC 21. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 4, n. 2, p. 20-38, jan./jun. 2007.

CAMPOS, Luiz Fernando de B. Metadados digitais: uma revisão bibliográfica da evolução e tendências por meio de categorias funcionais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, n. 23, 2007.

LOURENÇO, Cíntia A. Metadados: o grande desafio na organização da web. **Informação & Sociedade: Estudos**, João Pessoa, v. 17, n. 1, p. 65-72, jan./abr. 2007.

SMIT, Johanna W. Arquivologia/Biblioteconomia: interfaces das ciências da informação. **Informação & Informação**, Londrina, v.8, n. 1, jun./dez. 2003.



SMIT, Johanna W. O documento audiovisual ou a proximidade entre as 3 Marias.
Revista Brasileira de Biblioteconomia e Documentação, São Paulo, v. 26, n. 1/2,
p. 81-85, jan./jun. 1993.



ANÁLISE ESTATÍSTICA DO ACERVO DE RECURSOS EDUCACIONAIS EM SAÚDE

Soraya Alves Lacerda¹, Fernanda de Souza Monteiro², Aline Santos Jacob³

¹Universidade Aberta do SUS (UNA-SUS); ²Universidade de Brasília (UnB) / Faculdade de Ciência da Informação / Departamento de Biblioteconomia; ³Universidade Aberta do SUS (UNA-SUS)

Pôster: http://biredial.ufr.br/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/147/135

Resumo

Apresenta um estudo quantitativo da submissão, acesso e *download* de recursos educacionais no Acervo de Recursos Educacionais em Saúde (ARES). Este trabalho constitui-se a partir do uso do módulo de estatísticas do Dspace, acrescido de dados do PiWik, uma ferramenta open source de *web analytics*. Em sua contextualização, destaca-se o legado de conteúdos produzidos pelas iniciativas de ensino-aprendizagem para profissionais de saúde, financiadas em todo país pelos poderes públicos, tais como o Sistema Universidade Aberta do SUS (UNA-SUS). Com esses conteúdos disseminados no ARES, observou-se a necessidade de avaliar dados e estabelecer indicadores que orientassem ações capazes de ampliar a visibilidade desse repositório, bem como, aprimorar suas funcionalidades para atender o público-alvo. Isto justifica-se pelo fato do ARES ter completado um ano de atividades no primeiro semestre de 2013, com aproximadamente 600 recursos na época. A metodologia adotada foi a definição de um marco inicial, a configuração das ferramentas necessárias para medir e armazenar dados estatísticos e, posteriormente, a análise dos resultados de acordo com o contexto. Vale destacar, que o Sistema UNA-SUS é constituído pela Rede UNA-SUS, com Instituições de Ensino Superior (IE) que oferecem capacitação e formação continuada aos trabalhadores do SUS, por meio de cursos a distância; pela Plataforma Arouca, com informações cadastrais e acadêmicas desses profissionais; e o ARES. A alimentação do ARES se dá de forma federativa pela Rede, preservando e tornando pública sua produção. O monitoramento estatístico do repositório busca não só acompanhar as métricas do seu alcance e do seu uso, mas também incentivar o depósito pelas IE's, com base nos indicadores de visibilidade. Assim, observou-se quais números são mais relevantes no monitoramento do ARES, traçando um comparativo entre os números fornecidos pelo módulo de estatísticas do DSpace e pelo PiWik. As conclusões permitiram a realização de inferências que fundamentaram o planejamento da nova versão do repositório e o delineamento de estratégias para ampliar sua visibilidade.

Palavras-Chave: Repositórios; Métricas; Estatísticas de Acesso; Web Analytics



Referências

CHRISTOPHER, Joe. **How to track downloads in Google Analytics**. 2011. Disponível em: <<http://www.blastam.com/blog/index.php/2011/04/how-to-track-downloads-in-google-analytics/>>. Acesso em: 23 jul. 2013.

DURASPACE. **DSpace statistics**. Disponível em: <<https://wiki.duraspace.org/display/DSDOC3x/DSpace+Statistics>>. Acesso em: 10 jul. 2013.

FOSCARINI, Éverton Didoné; MARCHI, Alexandre Albino. **PiWik**: uma alternativa livre ao Google Analytics. 2009. Disponível em: <<http://www.lume.ufrgs.br/bitstream/handle/10183/16097/000696924.pdf?sequence=1&locale=en>>. Acesso em: 10 jul. 2013.

MODULE: Dspace statistics and Google Analytics. Disponível em: <<http://cadair.aber.ac.uk/dspace/bitstream/handle/2160/623/Module+-+DSpace+statistics+and+Google+Analytics.pdf;jsessionid=50FE55AF0201A243BA0E2BD3C105DB70?sequence=3>>. Acesso em: 10 jul. 2013.

SHEPPERD, Nick. **Using Google Anaaalytics statistics within DSpace**. 2012. Disponível em: <<http://ukcorr.blogspot.com.br/2012/03/using-google-analytics-statistics.html>>. Acesso em: 23 jul. 2013.



CARACTERIZAÇÃO DO PROFISSIONAL DA INFORMAÇÃO EM REPOSITÓRIOS DE ACESSO ABERTO NO BRASIL

Elis Gabriela Copa dos Santos, Ariadne Chloë Mary Furnival

¹Estudante de Biblioteconomia e Ciência da Informação da Universidade Federal de São Carlos. Estagiária da Secretaria de Educação à Distância da UFSCar); ²Professora adjunta da Universidade Federal de São Carlos

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/166/110

Resumo

Entre os muitos desafios enfrentados por profissionais da informação hoje, o tratamento da informação digital é, sem dúvida, um dos maiores. No tocante à informação científica, existem muitas questões. Entre elas está a sua disponibilização em repositórios de acesso aberto, muitos destes estabelecidos, gerenciados e mantidos em universidades pelo mundo todo (os chamados “repositórios institucionais” – RIs). Há, inclusive, um consenso em grande parte da literatura que trata o tema de repositórios digitais que a responsabilidade pela realização desse conjunto de atividades relacionadas ao repositório recai sobre as respectivas Bibliotecas Universitárias e seu quadro de pessoal formado na área de biblioteconomia, usualmente contando com o apoio contínuo e integrado de pessoal técnico dos setores de informática da universidade em questão. Mas é indiscutível que, na vida útil do repositório (pós-implementação), as tarefas principais envolvidas na alimentação, processamento, divulgação e avaliação do seu conteúdo, são atribuídas aos profissionais de informação, usualmente bibliotecários. Assim, a **questão da pesquisa** aqui proposta discute do papel desenvolvido pelo profissional da informação que atua por trás desses repositórios, focando nas competências que serão demandas desse profissional e a relação da atuação dele como mediador (ou não) no processo de depósito de itens no repositório. O **objetivo geral** deste trabalho é compreender e delinear o papel do profissional da informação (bibliotecário) no universo dos repositórios de acesso aberto brasileiros, bem como realizar pesquisa bibliográfica sobre a evolução do papel do bibliotecário e sua relação com a ampliação do alcance movimento do acesso aberto, sobre o início e fortalecimento da implantação de repositórios digitais em universidades e centros de pesquisa e sobre o binômio “Auto-arquivamento “vs.” Arquivamento mediado”. A **importância deste trabalho está justificada** no fato de que no Brasil ainda existem poucos trabalhos voltados à compreensão da relação entre os repositórios e os profissionais que atuam em sua gestão, bem como nas dificuldades encontradas neste processo e nas habilidades demandas e desenvolvidas por este profissional. O **método** adotado é de caráter exploratório, usando métodos qualitativos e quantitativos, já que engloba tanto a fase de levantamento, estudo e análise e síntese da literatura relevante ao estudo, quanto a fase empírica de levantamento de dados, composta pelo desenvolvimento e envio de questionários online aos gestores e demais envolvidos na gestão dos repositórios institucionais de universidades e centros de pesquisa brasileiros e a coleta e análise de 21 anúncios de emprego surgidos na Inglaterra e mais três em outros países entre 2012 e o primeiro



semestre de 2014 buscando por profissionais da informação para lidar com informação digital. Os **resultados esperados** são os de elaborar apontamentos e delineamentos iniciais para o *blueprint* (caracterização ideal) das competências necessárias para o profissional de informação visando traçar o perfil das competências desejáveis ao *Repository Manager* – o profissional de informação responsável pelo dia-a-dia do repositório, bem como sugerir possíveis recomendações relativas aos conteúdos necessários na formação desses profissionais aqui no Brasil.

Palavras-Chave: Profissional da informação; Repositórios Institucionais; Competências informacionais; Perfil profissional; Acesso Aberto



A VISÃO DO PESQUISADOR NO AUTOARQUIVAMENTO EM REPOSITÓRIOS INSTITUCIONAIS E A CONTRIBUIÇÃO DA INTERAÇÃO HUMANO-COMPUTADOR

Viviane Santos de Oliveira Veiga, Aline Silva, Rejane Machado, Denise Nacif Pimenta, Cícera Henrique da Silva, Maria da Conceição Carvalho, Luis Guilherme Gomes de Macena

Instituto de Comunicação e Informação Científica e tecnológica em Saúde. Fiocruz

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/170/109

Resumo

O Repositório Institucional é uma das ferramentas das Tecnologias de informação e comunicação (TICs) que estão alterando o processo de comunicação da ciência e seus impactos na sociedade. Objetiva principalmente melhorar a comunicação científica interna e externa à instituição; maximizar a acessibilidade, o uso e a visibilidade (Autor, 2009). Para o sucesso de um repositório, é preconizado o self-archiving, isto é, o depósito, pelo autor ou pessoa autorizada por ele, de um documento digital no sistema. A autonomia do autor na realização da tarefa de self-archiving de suas pesquisas científicas pode não ser exercida devido a diversos fatores, dentre eles, a usabilidade da interface. Pesquisas bibliográficas indicam a pouca frequência de estudos de usabilidade em repositórios institucionais envolvendo usuários na tarefa de self-archiving. Tarefa esta, fundamental no processo de mudança de comportamento e cultura dos pesquisadores na adesão ao repositório. Estes estudos também podem auxiliar a entender o modelo mental dos pesquisadores em cada área do conhecimento, uma vez que estudos apontam existirem diferenças entre a adesão ao autoarquivamento e a área do conhecimento e que há resistência por parte dos pesquisadores sênior a alterar o seu comportamento de publicação. ARCA, the Institutional Repository of the Fundação Oswaldo Cruz Fiocruz, linked to the Brazilian Ministry of Health, lançado em 2011, objetiva ser o principal instrumento de organização, preservação e disseminação da produção científica institucional e conta atualmente com 5.106 documentos digitais. A maioria destes foi depositada por profissionais de informação e não pelos pesquisadores. Visando então identificar possíveis dificuldades no autoarquivamento, foi desenvolvido um estudo de usabilidade do repositório ARCA com cinco pesquisadores da Fiocruz na realização da tarefa de self-archiving. Utilizou-se a avaliação cooperativa e aplicação de questionário pós teste. Os resultados obtidos demonstram que, apesar do reconhecimento dos benefícios proporcionados pelo uso do repositório institucional em termos de visibilidade da sua produção científica, algumas barreiras da interface enfrentadas pelos usuários podem contribuir para a não adesão ao repositório, enfraquecendo a proposta de melhoria da comunicação científica e a visibilidade da produção científica de uma instituição. Problemas de linguagem e ausência de feedback imediato na autorização para depósito. Em determinadas etapas do processo o excesso de informações desviou o foco na realização da tarefa, fazendo com que esta demande



mais tempo do que o necessário e muitas vezes induzindo ao erro. Por outro lado, a falta de informações em outros momentos traz dificuldades no processo e pode até causar a desistência do pesquisador em finalizar a tarefa. Propõe-se então possíveis soluções para os problemas de usabilidade identificados.

Palavras-Chave: Repositorio Institucional; Usabilidade; Autoarquivamento

Referências

LEITE, F.C.L.. *Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto*. Ibict, Brasília 2009.



EXPERIENCIAS Y APRENDIZAJES EN LA IMPLEMENTACIÓN DE UNA BIBLIOTECA DIGITAL PARA EL CENTRO DE INVESTIGACIÓN Y DOCUMENTACIÓN PARA LA ARTESANÍA – CENDAR: PRESERVACIÓN DE LA PRODUCCIÓN INTELECTUAL DE ARTESANÍAS DE COLOMBIA

Linamaría Pinzón Valencia
Bibliotecóloga en el Centro de Investigación para la Artesanía - Cendar

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/182/119

Resumen

El Centro de Investigación y Documentación para la Artesanía CENDAR, fue creado en el año de 1978 con el propósito de organizar la información del sector artesanal del país y la generación de políticas para la producción artesanal de Artesanías de Colombia. Desde su inicio, el centro de documentación ha buscado fomentar proyectos de investigación del sector artesanal y generar un ambiente de motivación para el avance de la investigación y el conocimiento, así como para la producción, acopio, clasificación y difusión de los recursos bibliográficos sobre la artesanía y el arte popular. Debido a la centralización de la información y documentación del acervo bibliográfico del CENDAR y de su única sede en Bogotá, en el año 2012 se puso en marcha el proyecto de “Investigación y Gestión del Conocimiento” realizado por la directora del Centro de documentación – Gladys Salazar Garcés, con el propósito de lograr descentralizar la información técnica que ha producido Artesanías de Colombia, elaborada principalmente por contratistas que se han vinculado a la entidad para la ejecución de proyectos en diferentes áreas y sitios diversos de la geografía colombiana, esto con el fin de ponerla a disposición de la comunidad interesada y del público general, de manera que puedan acceder a experiencias significativas que les permitan comprender el sector, ahondar en sus desarrollos y estimular la investigación en aras de convertir la artesanía en un sector que dé cabida a la innovación y logre la competitividad a nivel nacional e internacional. Artesanías de Colombia, como entidad rectora del Sector artesanal vinculada al Ministerio de Industria, Comercio y Turismo, produce una cantidad importante de información, al igual que muchos otros actores del sector, que constituye la materia prima del conocimiento que se produce sobre la artesanía y el sector artesanal, pero que infortunadamente no se trabaja con el fin exclusivo de producir conocimiento, de manera que éste sirva de base para la formalización del sector y la toma de decisiones en aras de alcanzar su fortalecimiento, desarrollo y competitividad. Por esta razón, se dio inicio al proceso de digitalización de los documentos institucionales, mediante la implementación de la Biblioteca Digital con el fin de brindar acceso y visibilidad a la documentación producida por Artesanías de Colombia, utilizando un software de código abierto, DSPACE utilizado para la administración de colecciones digitales bajo el uso de licencias de Creative Commons, esta actividad permitió dar cabida al fortalecimiento de los procesos de



catalogación de documentos digitales especializados sobre el sector artesanal. La biblioteca digital de Artesanías de Colombia, le permite a toda la comunidad desde diferentes lugares del país y del mundo, que consulten, descarguen y distribuyan los documentos en texto completo de manera abierta y gratuita con el fin de promover el acceso abierto a la producción científica e institucional, y enlazar la información y el conocimiento para fortalecer los procesos de investigación en el sector, de esta forma, el conocimiento institucional, especialmente de las personas que han trabajado con las comunidades de artesanos se podrán conocer públicamente y ser parte fundamental del rescate del trabajo del artesanado colombiano. Esta ponencia, quiere dar a conocer las experiencias y los aprendizajes del resultado de la implementación de la Biblioteca Digital en el Centro de Investigación y Documentación de Artesanías de Colombia, con miras a mejorar el acceso, la visibilidad, el impacto y la preservación de la documentación, no solo por su producción institucional sino del sector artesanal colombiano, que permita el desarrollo de nuevos conocimientos para el usufructo de la institución, la academia, las organizaciones de artesanos, los comercializadores y el sistema productivo en general, como fuente principal y materia prima de la actividad artesanal.

Palabras-Clave: Acceso abierto, Biblioteca Digital, Preservación Digital, Artesanías de Colombia

Citas

Misión. Artesanías de Colombia. [En línea]. Consultado en: www.artesantiasdecolombia.com.co

Salazar Garcés, María Gladys. (2012). Proyecto de Investigación y Gestión del Conocimiento. Artesanías de Colombia.



A (IN)DEFINIÇÃO DO BACKUP NO CONTEXTO DAS ESTRATÉGIAS DE PRESERVAÇÃO DIGITAL: UM PROCEDIMENTO ESTRUTURAL OU OPERACIONAL?

Henrique Machado dos Santos¹, Daniel Flores, Jorge Alberto Soares Cruz², Dhion Carlos Hedlund³, Evandro Anacleto Abreu Ferreira⁴

¹Acadêmico do Curso de Arquivologia, aluno especial e bolsista do Mestrado em Patrimônio Cultural da Universidade Federal de Santa Maria; ² Universidade Federal de Santa Maria (UFSM); ³ Fundação Universidade Federal do Rio Grande (FURG); ⁴ Acadêmico do curso de Arquivologia da UFSM

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/183/120

Resumo

Este trabalho aborda a contextualização teórico-conceitual dos procedimentos de *backup* frente a preservação digital. Com esta abordagem tem-se por objetivo verificar a viabilidade do *backup* como estratégia de preservação digital, identificado o seu nicho e sua necessidade de aplicação. A metodologia consiste na pesquisa bibliográfica realizada em fontes como: teses, diretrizes, livros e artigos. Sendo assim, a pesquisa se configura como uma revisão de literatura abordada de forma qualitativa. Primeiramente são definidos os conceitos de estratégias de preservação e procedimentos de *backup*, em um segundo momento é discutido a sua implementação com vistas na preservação de longo prazo, e por fim é definido a sua natureza frente a preservação digital. Dentre as definições de preservação digital, esta deverá garantir o acesso à informação autêntica (FERREIRA, 2006) e integra (CONARQ, 2004), a qual será interpretada por uma plataforma tecnológica no contexto futuro (FERREIRA, 2006; CONARQ, 2004). Isto dependerá da solução tecnológica proposta e dos seus respectivos custos (MÁRDERO ARELLANO, 2004), considerando o acesso contínuo ao conteúdo do documento digital (CONARQ, 2004). A preservação digital é composta por dois tipos de procedimentos: os de ordem estrutural que dizem respeito a definições de normas, adoção de padrões e a infraestrutura; e os de ordem operacional que são as atividades aplicadas para a preservação física, lógica e intelectual dos documentos digitais (THOMAZ, 2004; MÁRDERO ARELLANO, 2004). Neste momento questiona-se se a natureza do *backup*: ele é um procedimento estrutural ou operacional? Segundo Franch (2008) o *backup* consiste na replicação idêntica dos dados que integram um documento, esta é uma técnica comum em qualquer contexto tecnológico, que se tornou necessária devido as experiências de perdas em nível mundial. Franch define o *backup* como uma estratégia de preservação direcionada ao nível conceitual do objeto digital, evidenciando-o como um procedimento operacional. Innarelli (2012) por sua vez comenta que a política de *backup* é o primeiro passo para a preservação de documentos digitais, garantindo integridade, confiabilidade e a possível restauração dos dados. Em síntese, Innarelli realça a importância do *backup* assim como Franch, porém vai contra este ao definir o *backup* como uma política de segurança da informação, e conseqüentemente enquadrando-o como um procedimento de ordem estrutural. Para Fontana, Flores, Nora e Santos (2014) o *backup* é uma estratégia de preservação não consensual, embora se realizado em local distinto dos arquivos originais assegure que a informação esteja salva, ele não proporciona garantia de acesso contínuo, pois não contempla a atualização de

formatos de arquivo e versões de *software*. Além disto, poderá sofrer obsolescência do próprio suporte. O Interpares (2007) recomenda que o *backup* seja realizado sobre o sistema abrangente, isto inclui o sistema operacional, os *softwares* e os materiais digitais do sistema. Considerando as questões abordadas, percebe-se que as rotinas de *backup* são procedimentos fundamentais para desenvolver o plano de preservação digital, esta necessidade conceitua o *backup* como parte das políticas de preservação digital. Em resposta a questão levantada anteriormente, conclui-se que o backup é um procedimento de ordem estrutural.

Palavras-Chave: Preservação digital, Backup, Documentos digitais, Segurança da informação

Referências

CONSELHO NACIONAL DE ARQUIVOS – CONARQ (Brasil). Câmara Técnica de documentos eletrônicos. 2004. Carta para a Preservação do Patrimônio Arquivístico Digital. Rio de Janeiro: Arquivo Nacional, 2004. Disponível em: <<http://www.conarq.arquivonacional.gov.br/Media/publicacoes/cartapreservpatrimarqdigitalconarq2004.pdf>> Acesso em: 20 mai. 2014.

FERREIRA, Miguel. Introdução à preservação digital – Conceitos, estratégias e atuais consensos, Portugal: Escola de Engenharia da Universidade do Minho, 2006. Disponível em: <<https://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf>>. Acesso em: 02 abr. 2014.

FONTANA, Fabiana Fagundes; FLORES, Daniel; NORA, Fábila Dalla; SANTOS, Henrique Machado dos. Archivematica como ferramenta para acesso e preservação digital à longo prazo. *ÁGORA*, ISSN 0103-3557, Florianópolis, v. 24, n. 48, p. 62-82, 2014. Disponível em: <<http://agora.emnuvens.com.br/ra/article/view/457>>. Acesso em: 29 mai. 2014.

FRANCH, David Iglésia. La fotografia digital em los archivos – Qué es y como se trata. Ediciones TREA, 2008, España.

INNARELLI, Humberto Celeste, Instrumenta 2: Preservação de Documentos Digitais. Associação dos Arquivistas de São Paulo. 64p. São Paulo: ARQ-SP, 2012.

INTERPARES 2 PROJECT. Diretrizes do Produtor. A elaboração e a manutenção de materiais digitais: diretrizes para indivíduos. TEAM Brasil. Tradução: Arquivo Nacional e Câmara dos Deputados. 2002 – 2007. Disponível em: <http://www.interpares.org/ip2/display_file.cfm?doc=ip2_creator_guidelines_booklet--portuguese.pdf> Acesso em: 09 abr. 2014.

MÁRDERO ARELLANO, Miguel Ángel. Preservação de documentos digitais, *Ci. Inf.*, Brasília, v. 33, n. 2, p. 15-27, maio/ago. 2004. Disponível em: <<http://www.scielo.br/pdf/ci/v33n2/a02v33n2.pdf>> Acesso em: 25 mar. 2014.



THOMAZ, Kátia de Pádua. A preservação de documentos eletrônicos de caráter arquivístico: novos desafios, velhos problemas. 389f. Tese (Doutorado em Ciência da Informação). Escola de Ciência da Informação. Universidade Federal de Minas Gerais, 2004.



PRODUÇÕES ACADÊMICAS EM REPOSITÓRIOS INSTITUCIONAIS: A BUSCA RETROSPECTIVA DE AUTORIZAÇÕES PARA DISPONIBILIZAÇÃO EM ACESSO ABERTO

Diane Cátia Tomasi, Rubens Da Costa Silva Filho
Bibliotecários Biblioteca da Escola de Enfermagem - UFRGS

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/185/111

Resumo

Introdução: Pesquisa experimental desenvolvida na Biblioteca da Escola de Enfermagem da UFRGS visando a preservação e o aumento da visibilidade da coleção de Trabalhos de Conclusão de Curso de Graduação (TCC), período 2001-2005, através de sua disponibilização no repositório institucional LUME da UFRGS. A coleção impressa de TCCs ao ser disponibilizada no LUME, proporcionará considerável ganho de espaço físico no acervo, além de padronizar o suporte da coleção de TCCs, evitando também a duplicidade de suportes informacionais na biblioteca. Segundo Leite (2009), os repositórios digitais trazem muitos benefícios para as universidades entre eles o aumento da visibilidade, reputação e prestígio da instituição mantenedora. Hollós e Silva (2010), afirmam que a preservação documental caracteriza-se como uma atividade responsável pela gestão dos recursos dedicados a garantir a integridade física dos documentos, proporcionando o aumento de sua durabilidade e proporcionando o acesso às gerações presentes e futuras. **Materiais e Métodos:** Após mapeamento da coleção impressa de TCCs no período 2001-2005, disponíveis no acervo, e para que fosse possível a disponibilização eletrônica no LUME, foi necessário realizar a busca pelo contato (email) dos autores para solicitar a autorização dos mesmos para disponibilizar seu trabalho no repositório. A busca pelos contatos se deu através do cadastro do ex-aluno na UFRGS, não havendo sucesso nessa busca, foi realizada pesquisa no Google e Facebook. Através do Facebook foi enviada uma mensagem “in box” para o autor solicitando seu e-mail para posterior envio do formulário de autorização. Após retorno das mensagens pelo Facebook foram enviados e-mails formais para a solicitação das autorizações. **Resultados Parciais:** Como resultado parcial do estudo, dos 207 TCCs do período pesquisado, 90 autores foram localizados, 68 retornaram com a autorização de disponibilização no repositório, e 22 autores ainda não retornaram o e-mail. **Considerações parciais:** A avaliação preliminar do projeto é positiva, pois a biblioteca otimizou o espaço físico de seu acervo e proporcionou maior visibilidade aos TCCs, uma vez que os trabalhos que estavam disponíveis apenas em meio impresso no acervo da biblioteca foram digitalizado e estão disponíveis on-line. O trabalho pretende se estender às demais coleções que são Produção Intelectual (PI) da faculdade, como teses, dissertações e relatórios técnico-científicos.

Palavras-Chave: Repositórios institucionais. Biblioteca universitária. Avaliação de coleções



Referências

HOLLÓS, Adriana Cox; SILVA, Rubens R. G. Parâmetros de atuação do conservador na área da preservação documental. In: SILVA, Rubens Ribeiro Gonçalves et al (Org.). **Cultura, representação e informação digitais**. Salvador: EDUFBA, 2010. p. 17-28.

LEITE, Fernando César Lima. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto. Brasília, DF: Ibict, 2009. Disponível em: <<http://livroaberto.ibict.br/bitstream/1/775/4/Como%20gerenciar%20e%20ampliar%20a%20visibilidade%20da%20informa%C3%A7%C3%A3o%20cient%C3%ADfica%20brasileira.pdf>>. Acesso em: 28 maio 2014.

INSTITUTIONAL REPOSITORIES AND THE DUPLICATE OF ACADEMIC PRODUCTIONS IN UNIVERSITY LIBRARIES

Abstract

Introduction: Experimental research development at the Library of the School of Nursing at UFRGS aiming at preserving and increasing the visibility of the collection of Work Conclusion of Graduate Course (TCC), 2001-2005, through its institutional repository available on LUME UFRGS. The printed collection of TCCs to be available on LUME, provide considerable gain physical space in the collection, in addition to standardizing support the collection of TCCs also avoiding duplication of informational media in the library. According to Leite (2009), digital repositories bring many benefits to universities including increased visibility, reputation and prestige of the supporting institution. Hollós and Silva (2010), claim that the document preservation is characterized as a responsible for managing the resources dedicated to ensure the physical integrity of documents activity, providing increased durability and providing access to present and future generations. **Materials and Methods:** After mapping the printed collection of TCCs in 2001-2005, available in the collection, and to make possible the electronic availability on LUME, it was necessary to search for the contact (email) authors to request their authorization to deliver their work to the repository. The search for the contacts occurred through the register by former student at UFRGS, not having success in this search, research was conducted on Google and Facebook. Through Facebook was sent a message "in box" to the author requesting your email for later sending the authorization form. After returning messages by Facebook formal emails were sent to request authorizations. **Partial Results:** As a partial result of the study, of 207 TCCs of the research period, 90 authors were located, 68 returned with the authorization available in the repository, and 22 authors have not yet returned the email. **Partial Considerations:** A preliminary assessment of the project is positive, as the library optimized the physical



space of their collections and provided greater visibility to TCCs, since the works that were available only in print in the library collection were digitized and are available online. The work aims to extend to other collections that are Intellectual Production (IP) college, such as theses, dissertations, and technical and scientific reports.

Keywords: Institutional repositories. University library. Collections assessment



ACESSO DEMOCRÁTICO AO CONHECIMENTO: A INTEROPERABILIDADE DE RECURSOS EDUCACIONAIS DA BIBLIOTECA VIRTUAL DO NESCON E O REPOSITÓRIO ARES DA UNA-SUS

*Jacqueline Pawlowski Oliveira, Eliane Pawlowski Oliveira Araújo, Edison José
Correa, Maria Rizioneide Negreiros Araújo*
Universidade Federal de Minas Gerais

Pôster: http://biredial.ufr.br/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/190/127

Resumo

Informação é insumo fundamental para o desenvolvimento da ciência e manter-se atualizado e bem informado tem sido um grande desafio, principalmente quando se refere a área da saúde. Atualmente, diversas instituições compartilham informações buscando cooperação e interoperação com o objetivo de maximizar oportunidades de troca e reuso de informações, bem como contribuir com aperfeiçoamento e capacitação de profissionais. Segundo Miller (2000), interoperável – entendido como a capacidade de operar em conjunto - é um adjetivo cujo conceito é cada vez mais utilizado na gestão da informação. Esta capacidade é conhecida na literatura como interoperabilidade que apresenta a habilidade de transferir e utilizar informações de maneira uniforme e eficiente entre várias organizações e sistemas de informação. Em atenção a esta premissa, a Universidade Federal de Minas Gerais (UFMG), instituição participante da Rede do Sistema Universidade Aberta do Sistema Único de Saúde (UNA-SUS), desde 2012 vem compartilhando seu acervo de recursos educacionais em saúde. Este acervo, produzido pela equipe do Núcleo de Educação em Saúde Coletiva (Nescon) da Faculdade de Medicina da UFMG, no âmbito do Curso de Especialização em Estratégia Saúde da Família (CEESF), contempla materiais educacionais, de formatos diversificados (módulos de curso, animação, vídeo, imagem, texto, etc.) que podem ser utilizados pelos profissionais da saúde para autoinstrução ou por instituições interessadas em oferecer cursos na área da saúde. O objetivo deste trabalho é apresentar o projeto de implantação do acervo do Nescon no repositório do Acervo de Recursos Educacionais em Saúde (ARES) da UNA-SUS, bem como seu desenvolvimento e resultados obtidos até o momento. Para que o acervo do CEESF pudesse integrar-se ao ARES foi necessário definir um plano de ações e trabalhar em etapas diferenciadas pré-determinadas pela equipe da UNA-SUS. Completando dois anos de atividades no primeiro semestre de 2014, e com uma coleção de aproximadamente 1000 recursos educacionais, a UNA-SUS vem disseminando o conhecimento produzido por suas instituições parceiras e buscando dar um maior alcance ao ARES. No período de 1 ano e meio, o acervo do Nescon/UFMG incorporado ao ARES obteve um total de 8.719 acessos e 2.475 downloads de seu material compartilhado o que é um valor significativo considerando a quantidade de material disponibilizado no repositório. Atualmente, o Nescon está com 90% de seu acervo inserido no repositório ARES e vem trabalhando para compartilhar, na totalidade, toda sua produção. Espera-se que a



Rede UNA-SUS, por intermédio de suas cooperantes, constitua um acervo a partir de uma cultura de colaboração, levando a um incremento da quantidade e qualidade de sua coleção. A redução do desperdício de recursos com materiais e ofertas redundantes ocorrerá conseqüentemente, pois todas as instituições terão acesso a todos os recursos educacionais produzidos pela Rede. Muito já se avançou em termos de cooperação no âmbito das atividades próprias à Biblioteconomia e Ciência da Informação. É percebida, porém, a necessidade crescente de se aglutinarem esforços para que se otimize a eficiência, de forma interoperável, na área da saúde com foco no profissional desse campo do saber, especialmente no que toca à sua formação continuada.

Palavras-Chave: Repositório institucional; interoperabilidade; Recursos educacionais em saúde

Referências

BRASIL. Ministério da Saúde. Universidade Aberta do Sistema Único de Saúde. **Política do acervo de recursos educacionais em saúde**. Brasília: UNASUS, 2011.

MILLER, P. Interoperability: what is it and why should I want it? **Ariadne**. n. 24, jun. 2000. Disponível em: <<http://www.ariadne.ac.uk/issue24/interoperability/>>. Acesso em: 12 jan. 2014.

UNIVERSIDADE aberta do SUS. Acervo de Recursos Educacionais em Saúde – ARES. 2011. Disponível em: <<http://ares.unasus.gov.br/acervo/>>. Acesso em: 13 jan. 2014.



REPOSITÓRIO INSTITUCIONAL DA UNIVERSIDADE FEDERAL DO RIO GRANDE - FURG E O LIVRE ACESSO A PRODUÇÃO TÉCNICO-CIENTÍFICA DA INSTITUIÇÃO

Cristiane Castro Soares¹, Gabriela Silva da Rosa¹, Paula Porto Gauterio¹, Maria Helena Machado de Moraes², Angélica Conceição Dias Miranda²

¹Graduanda do curso de Biblioteconomia da Universidade Federal de Rio Grande – FURG;

²Professora do Curso de Biblioteconomia da Universidade Federal de Rio Grande - FURG

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/193/112

Resumo

Os Repositórios Institucionais (RIs) surgem com o objetivo de disponibilizar a produção intelectual da instituição de forma livre (sem custos) para acesso, consulta, pesquisa e download. O presente trabalho tem como objetivo ressaltar a importância do acesso livre a informação e apresentar o Repositório Institucional da Universidade Federal do Rio Grande – RI FURG. Como metodologia discute-se a abordagem qualitativa, uma vez que se apresenta a estrutura do RI FURG, a equipe e o passo a passo para o povoamento do mesmo. O RI FURG conta, até o momento, com 3828 produções distribuídas em 20 comunidades, assim como aumento da visibilidade da instituição, e resalta a importância do acesso a toda informação. Compreende-se que o RI FURG apresenta políticas voltadas para o processo de identificação, armazenamento, preservação, recuperação e disseminação da pesquisa científica dos servidores da instituição. Proporcionam assim, maior visibilidade para o pesquisador e para a instituição. Pode-se comentar que a universidade, a partir da implementação do seu repositório, contribui para o livre acesso a informação científica, apoiando a visibilidade e o crescimento da ciência.

Referências

RODRIGUES, E. et al. Repositório: criação e desenvolvimento do repositório institucional da universidade do Minho. Universidade do Minho, Braga (Portugal), 2004. Disponível em: <http://hdl.handle.net/1822/422>. Acesso em: 18 fev. 2014.



POLÍTICAS E PLANOS DE PRESERVAÇÃO DE OBJETOS DIGITAIS NAS INSTITUIÇÕES PÚBLICAS FEDERAIS DE ENSINO SUPERIOR: PANORAMA NACIONAL

Sonia Cruz-Riascos¹, Laura Vilela Rodrigues Rezende², Lais Pereira de Oliveira³
¹Profa. Dra. Universidade Federal de Pernambuco (UFPE), Centro de Artes e Comunicação; ²Profa. Dra. Universidade Federal de Goiás (UFG), Faculdade de Informação e Comunicação (FIC); ³Profa. Msc. Universidade Federal de Goiás (UFG), Faculdade de Informação e Comunicação (FIC);

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/194/121

Resumo

O presente trabalho trata da preservação de objetos digitais no contexto acadêmico de instituições públicas federais de ensino superior no país. Observa-se a proliferação de repositórios institucionais voltados para as publicações científicas, contudo o mesmo não vem ocorrendo, na mesma intensidade e preocupação, com a preservação digital e a custódia de recursos educacionais abertos. O problema abordado, portanto, recai sobre a diminuta incidência de iniciativas quanto à disponibilização de políticas e planos de preservação de repositórios para objetos digitais produzidos no meio acadêmico. Como objetivo maior deste estudo se teve a realização de mapeamento de ações para instituição de políticas para os repositórios desses recursos digitais, bem como de atividades para a preservação deles. Trata-se de pesquisa do tipo quantitativa, com abordagem descritiva, que teve como instrumento de coleta de dados o levantamento de dados ao nível nacional nos sites das universidades federais. A análise de dados aponta medidas empreendidas no cenário nacional. Comenta-se sobre perspectivas futuras para avanços no campo da preservação digital. Como resultados pretendidos, o aumento da importância da preservação e da visibilidade dos recursos educacionais abertos para a melhoria dos processos de ensino, aprendizagem e pesquisa no âmbito universitário. Conclui-se com reflexões relevantes sobre a preservação de objetos digitais e sua custódia no seio acadêmico nacional.

Palavras-Chave: Acesso aberto, Preservação digital, Política de preservação, Plano de preservação, Objetos digitais, Recursos Educacionais Abertos

Referências

BEAGRIE, Neil Beagrie; SEMPLE, Najla; WILLIAMS, Peter; WRIGHT, Richard. **Digital Preservation Policies Study: Part I Final Report October 2008**. London, JISC; Charles Beagrie Limited, copyright HEFCE 2008. 60 p.



COLEÇÃO DIGITAL DE JORNAIS DO MUSEU REPUBLICANO “CONVENÇÃO DE ITU” (MRCI-MP/USP): TRANSPOSIÇÃO DE SUPORTE PARA PRESERVAÇÃO E ACESSIBILIDADE

Márcia Medeiros de Carvalho Mendo¹, Maria Cristina Monteiro Tasca², Alline de Sousa³, José Renato Margarido Galvão², Rosemary Mendonça Martins Fernandes¹

¹Museu Paulista; ²Museu Republicano “Convenção de Itu”; ³Universidade de São Paulo

Pôster; http://biredial.ufr.br/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/196/122

Resumo

Relata o processo de digitalização da Coleção de Jornais Ituanos dos Séculos XIX e XX, composta de 11 títulos com 5.300 exemplares, depositados na biblioteca do Museu Republicano Convenção de Itu, que em parceria com a Biblioteca do Museu Paulista e o Departamento Técnico do SIBi/USP, objetivou preservar e tornar acessível o conteúdo dos jornais, referente aos campos da História Política, Social, Econômica e Jurídica do Brasil no período da configuração do regime republicano. A digitalização foi realizada ao longo de seis meses pela equipe após treinamento especializado, no Laboratório de Digitalização e Preservação Digital da USP, através da máquina SkyView, equipamento de ponta adquirido junto à empresa Kirtas Technologies (EUA), com uma câmera de 21.1 megapixel e sistema de achatamento por vácuo. As capturas geraram arquivos em formato TIFF 24bits em alta resolução, armazenados em servidores locais e tratadas no Software BSE para ajuste dos parâmetros de brilho e contraste, remoção de manchas e cor de fundo e demais correções necessárias. Em seguida, foram gerados arquivos de PDF completos, personalizados com a tecnologia de reconhecimento de caracteres (OCR – optical character recognition), obtendo-se um arquivo PDF pesquisável da obra que aprimora a busca do usuário. Os registros finais foram indexados e serão disponibilizados na Biblioteca Digital de Obras Raras, Especiais e Documentação Histórica da USP, disponível em <<http://www.obrasraras.usp.br>>, de consulta livre e gratuita, por meio da qual é possível fazer o download do PDF completo ou ler a obra online.

Palavras-Chave: Digitalização; Obras raras; Biblioteca digital; Preservação de acervos museológicos



Referências

ANDRADE, Ricardo; BORGES, Jussara; JAMBEIRO, Othon. Digitalizando a memória de Salvador: nossos presente e passado têm futuro? *Perspectiva em Ciência da Informação*, Belo Horizonte, v.11, n.2, p. 243-254. Maio/ago. 2006.

BARBUY, Heloisa; TASCA, Maria Cristina Monteiro. Coleção Digital de Jornais do Museu Republicano “Convenção de Itu” (MRCI/USP). São Paulo: DT-SIBI/USP, 2014. Disponível em: <http://obrasraras.sibi.usp.br/?p=41>. Acesso em: 06 maio 2014.

FREIRE, Eliane Oliveira de Lima; AZEVEDO, Rodolfo Almeida de. Salvaguarda do acervo documental da Reserva Florestal Ducke – Manaus-AM: registros e uma experiência. *Ciência da Informação*, Brasília, v. 39, n.2, p. 101-114, maio/ago. 2010.

GREENHALGH, Raphael Diego. Digitalização de obras raras: algumas considerações. *Perspectivas em Ciência da Informação*, v. 16, n.3, p. 159-167, jul./set. 2011.

MOREIRA, Alexandre et al. Digitalização de manuscritos históricos: a experiência da Casa Setecentista de Mariana. *Ciência da Informação*, Brasília, v. 36, n. 3, p. 89-98, set./dez. 2007.

POSSI, Maurilio de Araújo et. al. Ambiente para busca e visualizações de documentos históricos na web. *Perspectivas em Ciência da Informação*, v.16, n.3, p. 168-180, jul./set. 2011.

SANTOS, Vilma Moreira dos et al. A Coleção Linhares em meio digital. *Varia História*, Belo Horizonte, v.27, n. 46, p. 735-750. Jul./dez. 2011.

REIFSCHMEIDER, Oto Dias B. A. Importância do acesso às obras raras. *Revista Ibero-Americana de Ciência da Informação*, v.1, n.1, jan. / abr. 2008. Disponível em: <http://seer.bce.unb.br/index.php/RICI/article/view/1544/2330>. Acesso em: 06 maio 2014.



REDE CARINIANA: ATUALIDADES E PERSPECTIVAS

*Valéria Gameleira da Mota, Fernanda Weschenfelder
Instituto Brasileiro de Ciência e Tecnologia*

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/200/123

Resumo

O presente trabalho descreve os resultados obtidos pela Rede Brasileira de Serviços de Preservação Digital - Cariniana após um ano de seu início em fevereiro de 2013 quando o IBICT com o financiamento da FINEP aderiu à ferramenta e o Suporte do LOCKSS Technical Support Service (LTSS). O projeto inicialmente foi dividido em duas etapas, na primeira a equipe da Rede Cariniana organizou e planejou um serviço de preservação digital de periódicos eletrônicos convidando instituições públicas de ensino superior de diferentes estados que possuíam publicações digitais de acesso livre (OJS/SEER). Foram elaborados guias e manuais para auxiliar os parceiros e a elaboração de uma política da Rede com o objetivo de garantir os direitos e deveres das instituições envolvidas no projeto. A segunda etapa foi dedicada à expansão da Rede convidando outras instituições, formalização das parcerias através de acordos de cooperação técnica e a inclusão dos títulos brasileiros na lista global de publicações preservadas na Rede. O objetivo também é de ampliar os tipos de documento a serem preservados, incluindo os títulos do Portal Livro Aberto e das Teses e Dissertações da BDTD. A Rede Cariniana realizou neste período três encontros nacionais de parceiros integrais e um Seminário Internacional de Preservação Digital, que teve como objetivo de reunir profissionais dedicados à preservação digital. O objetivo da Rede Cariniana não é somente o de preservar documentos digitais, mas também de oferecer serviços que auxiliem os projetos que estejam relacionados à Preservação Digital, como a elaboração de cursos técnicos de digitalização, políticas e projetos, soluções tecnológicas, curadoria digital, arquivística e plataforma LOCKSS com o auxílio de colaboradores e especialistas, e a criação de onze grupos de pesquisa todos com assuntos relacionados à preservação digital, no qual irão auxiliar no desenvolvimento de serviços e produtos na área. Estão sendo agendadas reuniões com instituições de vários países como: Canadá, México, Costa Rica, Colômbia e Uruguai com o interesse de participar da Rede Cariniana.

Palavras-Chave: preservação digital, LOCKSS, cursos técnicos, periódicos eletrônicos



Abstract

This current work is to describe the results obtained by Brazilian Network of Digital Preservation Services - Cariniana after a year of its initialization in February of 2013 when the IBICT with FINEP funding and the support of LOCKSS technical Support Services (LTSS). The project was initially divided in two stages: in the first one Rede Cariniana team organized and planned a service of digital preservation of electronic periodicals to invite the superior education public organizations of different states which have free access digital publications (OJS/SEER). Guides and manuals were created in order to help the Cariniana network public universities, and the elaboration of policy of its network to guarantee the rights and obligation of the institutions involved in the project. The second stage was focused in the expansion of the network inviting the other institutions and the official partners formalization through technical cooperation agreements and the inclusion of Brazilian titles in the global list of publication preserved in the network. The objective is also to amplify the documentation types to be preserved, including the titles from "Portal Livro Aberto" and the theses and dissertations da BDTD. The Cariniana network realized during this period three national meetings of full time partners and one digital preservation International seminar that had the goal of gathering professionals dedicated to digital preservation. Cariniana's goal is not only to preserve documents but also to offer services to aide in projects related to digitalization such as elaborating technical classes in digitalization, policies and projects, technological solutions, digital collection, logistics in archive science, and a LOKSS platform with the help collaborators and specialists. The network also aim in creating eleven research groups all related to the digital preservation where they will assist in the development of services and products in this field. The Cariniana network is scheduling meetings with institutions of various countries like Canada, Mexico, Costa Rica, Colombia and Uruguay who have great amount of interest to participate in the network.



A BIBLIOTECA DIGITAL DO INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS COMO INSTRUMENTO DE PROMOÇÃO DO ACESSO LIVRE À PRODUÇÃO CIENTÍFICA NA ÁREA ESPACIAL

Clayton Martins Pereira, Marciana Leite Ribeiro, Jean Gerald Francis Banon
Instituto Nacional de Pesquisas Espaciais - INPE

Pôster: http://biredial.ufr.br/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/201/113

Resumo

A Iniciativa de Acesso Aberto (OAI) é um movimento liderado por cientistas de vários países, que tem por objetivo defender o acesso livre a toda produção científica disponibilizada na internet, de forma que qualquer cidadão possa ler, baixar e distribuir livremente seu conteúdo, democratizando assim o acesso à informação científica. Uma das estratégias apontadas por este movimento para se alcançar tal objetivo é por meio da criação de Repositórios Institucionais nas instituições de pesquisa. No Brasil, o Instituto Nacional de Pesquisas Espaciais (INPE) é o pioneiro entre as unidades de pesquisa do Ministério da Ciência, Tecnologia e Inovação (MCTI) ao criar um Repositório Institucional, também conhecido como “Biblioteca Digital da Memória Científica do INPE (BDMCI)”, com a finalidade de preservar e disseminar sua produção científica. Este trabalho aponta os desafios e oportunidades encontrados pelo Repositório Institucional do INPE na busca pela promoção do acesso livre à produção científica na área espacial. Para isso foi realizado um estudo que consistiu em verificar e contabilizar as permissões de acesso de cada produção científica depositada neste repositório, em especial a de artigos publicados em periódicos e em anais de congressos, considerando as políticas de autoarquivamento dos respectivos editores e organizadores. Os resultados obtidos mostraram que grande parte da produção científica publicada internamente (teses, dissertações, relatórios de pesquisa, *ePrints*, entre outros) e em anais de congressos está disponível em acesso livre, podendo ser acessada por qualquer cidadão a partir do portal da BDMCI[1]. No caso da produção científica publicada em periódicos, foi verificado que somente metade desta está disponível em acesso livre, por conta das políticas altamente restritivas ou de embargo das editoras comerciais. A partir da consolidação desses resultados foi possível apurar que 76% da produção científica depositada no Repositório Institucional do INPE está disponível em acesso livre. Isto permite concluir que o Repositório Institucional do INPE cumpre com sua meta de disponibilizar em acesso livre pelo menos 70% da produção científica nele depositada. No entanto, a exigência de manutenção e elevação das notas de avaliação dos programas de pós-graduação do instituto levou os pesquisadores a priorizar a publicação de artigos em periódicos, o que diminuiu drasticamente a quantidade de artigos publicados em anais de congressos. Para minimizar os impactos decorrentes, o INPE vem incentivando a produção e depósito de publicações internas, em especial os *ePrints* e os relatórios de pesquisa, bem como solicitado o depósito preferencial de versões “rascunho final” dos artigos



publicados em periódicos nos quais as respectivas editoras permitam sua disponibilização em acesso livre nos Repositórios Institucionais.

Palavras-Chave: Acesso Aberto, Repositórios Institucionais, Gestão da produção Científica

Referências

LEITE, F. C. L. Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto. Brasília: IBICT, 2009.

BANON, G. J. F.; RIBEIRO, M. L.; PEREIRA, C. M. Indicador de presença na web do acervo de um repositório institucional. In: CONFERÊNCIA LUSO-BRASILEIRA DE ACESSO ABERTO, 4. (CONFOA), 2013, São Paulo. **Pôsteres...** 2013.



PANORAMA DA PRODUÇÃO CIENTÍFICA BRASILEIRA EM MATEMÁTICA NOS REPOSITÓRIOS DE ACESSO LIVRE

Daniela Gralha de Caneda Queiroz¹, Ana Maria Mielniczuk de Moura²

¹Mestranda em Comunicação e Informação pela UFRGS; ²Professora no Departamento de Ciências da Informação da UFRGS e professora colaboradora no Programa de Pós Graduação em Comunicação e Informação (PPGCOM/UFRGS)

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/205/131

Resumo

A produção acadêmica na área da Matemática, no Brasil, ainda não atingiu os patamares de produção dos países desenvolvidos. Uma forma de mensuração disso são os repositórios digitais em ambientes abertos. Estas bases de dados são, atualmente, a melhor alternativa sem custo para a disseminação de ideias na comunidade científica dentro do ambiente virtual proporcionado pela rede mundial de computadores. O Brasil apareceu muito bem posicionado no ranking de revistas científicas de acesso livre em 2013, ocupando a segunda posição, com 923 revistas, apenas atrás dos Estados Unidos, com 1.249, conforme o Directory of Open Access Journals (2014). Entretanto, os resultados da produção matemática do país apresentam-se muito baixos e discute-se aqui os motivos desta baixa produção. O objetivo desse estudo, portanto, é apresentar um panorama da produção científica brasileira em Matemática nos repositórios de acesso livre, para que se possa estabelecer a relação direta entre os índices desejados e os atuais e os reflexos desta comparação no cenário acadêmico. É uma pesquisa do tipo bibliométrica, descritiva, sob forma de levantamento, com abordagem quantitativa. Seu corpus é composto pelos artigos de periódicos da área de Matemática produzidos por autores brasileiros e disponibilizados nos repositórios de acesso aberto Diretório Luso-Brasileira de Periódicos e Repositórios de Acesso Aberto, *Directory of Open Access Journals*, *Directory of Open Access Repositories* e *Directory of Open Access Repositories*. Os dados coletados referentes a autoria; vínculo institucional; nome e nacionalidade do periódico e ano de publicação do artigo foram tabulados no programa Microsoft Office Excel 2007, para a geração de tabelas e gráficos. O estudo encontra-se em andamento, mas análise preliminar dos dados indica incipiente participação brasileira na área de Matemática nos repositórios de acesso aberto, com provável destaque para a subárea de Ensino de Matemática. Considera-se que a baixa produção nessa área tem raízes mais profundas. Em 2009, avaliação feita pelo Programa Internacional de Avaliação de Alunos - PISA (2014), realizado pela Organização pela Cooperação e Desenvolvimento Econômico (OCDE), posicionou os alunos brasileiros no 54º lugar. As deficiências no ensino da matemática no ensino fundamental resultam em uma menor procura por essa área e, por consequência, podem estar refletindo em baixos índices de produção acadêmica nesse segmento. Os indicadores de sala de aula têm peso decisivo para a evolução dos indicadores econômicos de um país, especialmente na área da Matemática, e o crescimento do país tem relação direta com o ensino da Matemática e a produção científica, sendo o ranking da OCDE um reflexo direto desta correlação (PADUAN, 2014).

Palavras-Chave: Acesso aberto. Produção científica. Matemática./Open access. Scientific research. Mathematics



Referências

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira Pisa. **Pisa**. 2014. Disponível em: <<http://portal.inep.gov.br/pisa-programa-internacional-de-avaliacao-de-alunos>>. Acesso em: 25 maio 2014.

PADUAN, Roberta. O Impa dá um reforço no ensino de matemática no Brasil. **Exame**, São Paulo, n. 1058, 07 fev. 2014. Disponível em: <<http://exame.abril.com.br/revista-exame/edicoes/1058/noticias/um-reforco-no-ensino-de-matematica?page=3>>. Acesso em: 25 maio 2014.

OPEN ACCESS JOURNAL. 2014. Disponível em: <<http://doaj.org/>>. Acesso em: 25 maio 2014.

Abstract

In Brazil, scientific research in mathematics is not yet in the same level than developed countries. One way to evaluate this is the open access digital repositories. These databases are, nowadays, the best free of charge alternative to share ideas in the scientific community at World Wide Web. Brazil was well placed at scientific open journals world ranking in 2013, in 2nd with 923 journals, only behind USA with 1.249, according Directory of Open Journals (2014). However, the results of mathematics scientific research are very low, and here we discuss the reasons for this. The purpose, therefore, is to present Brazil's open repositories scientific research scenario in mathematics, so we can debate the difference between present and ideal levels and its reflexes in Brazilian scientific community. It's a bibliometrics research, with a quantitative approach. It contains the journal articles made by Brazilian researchers in mathematics available in Luso-Brasileira Directory of Journals and Repositories of Open Access, Directory of Open Access Journals, Directory of Open Access Repositories e Directory of Open Access Repositories. All data concerning authors, institution affiliation, name and country of the journal were gathered in Microsoft Office Excel 2007 to generate graphics. The research is not completed, but preliminary analysis shows little Brazilian mathematics research in open repositories, and most of them are in the field of Mathematics Basic Learning. These low results have deep reasons. In 2009, Programme for International Student Assessment – PISA (2014) from Organisation for Economic Co-operation and Development (OCDE) placed Brazilian students in 54th place. Difficulties in Mathematics Basic Learning causes low choice for this field in college and university and this may reflects the low levels of scientific research in this field. Basic educational levels have direct impact in the economy development and the country growth is a result of Mathematics Basic Learning and scientific research. The OCDE ranking is just a reflex of this association (PADUAN, 2014).



AUTORIA DE DADOS CIENTÍFICOS: PENSANDO A ANÁLISE A PARTIR DA TEORIA ATOR-REDE

Jackson da Silva Medeiros¹, Sônia Elisa Caregnato²

¹Doutorando em Comunicação e Informação na Universidade Federal do Rio Grande do Sul; ²Doutora em Information Studies pela University of Sheffield. Professora do Departamento de Ciências da Informação e do Programa de Pós-Graduação em Comunicação e Informação na Universidade Federal do Rio Grande do Sul

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/208/117

Resumo

A produção de dados científicos tem provocado mudanças no modo em como a ciência e a comunicação científica se apresentam. Esses dados compõem a e-Science, isto é, uma infraestrutura que possibilita a coleta dos dados, sua curadoria, compartilhamento, acesso e (re)uso. No entanto, as questões que ela enfrenta vão além de seu âmbito tecnológico, uma vez que ela não nasce fora de um contexto social e político, fazendo com que esses fenômenos tenham concepção complexa. Há, neste sentido, a percepção de delineamento do que significa a questão da autoria para o compartilhamento de dados científicos, refletindo, dessa maneira, um recorte que busca compreender a questão da autoria no âmbito dos dados digitais de pesquisa, buscando o aporte na Teoria Ator-Rede, a qual propõe um estudo focado na não delimitação de espaços e/ou atores (humanos ou não) para a investigação de associações em vez de preestabelecer categorias contextuais. Para que isto seja realizado é necessário que seja utilizada uma proposta sociotécnica, isto é, observar o processo de construção dessa rede com a imperiosa possibilidade de analisar de forma recursiva tanto humanos quanto objetos, bem como processos, atos, ações e escolhas, permitindo que sejam enfatizados os movimentos, as mudanças da rede. A utilização da Teoria Ator-Rede ocorre porque em situações que a inovação é cerne da questão proposta – como os dados científicos digitais e a autoria destes – os limites são ainda inconclusivos, necessitando que não sejam delimitados inicialmente, nem os atores que devem ser analisados e combinados para que as associações sejam estabelecidas, nem suas posições na estrutura do campo ao estabelecerem relações diversas. Há de reconhecermos a dificuldade que vem a ser o trabalho com a questão autoral e, em um primeiro olhar, dois pontos parecem merecer destaque: o primeiro se refere à própria complexidade na noção de autoria, investigação que diversos pesquisadores vêm buscando resultados. Outro ponto se refere a entender esta questão quando ligada aos dados científicos digitais e seu compartilhamento, uma vez que a pouca quantidade de trabalhos que buscam compreender o tema está ligada, essencialmente, à pesquisa em documentos bibliográficos, através, principalmente, da análise de discurso. A aplicabilidade desta proposta é apresentada de maneira a compreender o desenrolar do que consiste a rede de atores humanos e não humanos, descrevendo/narrando a representação do que existe culturalmente em uma perspectiva de rede. Parece razoável considerar que, para o universo estudado, o processo em análise precisa criar/encontrar as teias autônomas que se formam, reconhecendo seus atores e suas posições, bem como sua própria estrutura orgânica.

Palavras-Chave: Autoria, Dados digitais de pesquisa, Teoria Ator-Rede

Referencias

CHARTIER, R. **A aventura do livro: do leitor ao navegador: conversações com Jean Lebrun.** São Paulo: UNESP; Imprensa Oficial do Estado, 1998. 159 p.

CHARTIER, R. **Os desafios da escrita.** São Paulo: UNESP, 2002. 144 p.



TRATAMENTO DA INFORMAÇÃO E SUBJETIVIDADE: USANDO AS DIMENSÕES SIMBÓLICAS E AFETIVAS NA COMPREENSÃO DA CONSTRUÇÃO DOS REFERENCIAIS DE BUSCA EM SISTEMAS DE GESTÃO DE BIBLIOTECAS

Eliane Pawlowski de Oliveira Araujo¹, Claudio Paixão Anastácio de Paula²

¹Doutoranda em Ciência da Informação da UFMG; ²Professor Doutor do Departamento de Teoria e Gestão da Informação. Escola de Ciência da Informação. Universidade Federal de Minas Gerais

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/228/132

Resumo

É interessante observar como o trato com a informação tem se alterado, principalmente nas décadas de 1990 e 2000. De uma preocupação inicial com a disseminação da informação de modo a permitir ao usuário ter disponível todo o conteúdo necessário para satisfazer suas necessidades passou-se ao objetivo de procurar propiciar os meios mais adequados para que o usuário possa recuperar a informação que atenda sua demanda de forma específica. Todo esse cenário decorre do grande volume de informação disponível para a sociedade a partir do desenvolvimento das tecnologias de informação e comunicação, o que tornou mais complexa as atividades relacionadas ao tratamento da informação, principalmente no que diz respeito a recuperação. Garcia (2007) aponta como a recuperação da informação tem sido preocupação antiga, citando o “mundaneum” idealizado por Paul Otlet no Século XIX, que tinha como objetivo tornar acessível o conhecimento àqueles que dele necessitavam, assim como o “memex”, desenvolvido por Vannevar Bush na década de 1940, que visava facilitar a recuperação da informação. No âmbito das Bibliotecas a recuperação da informação é um dos aspectos mais importantes decorrentes do tratamento da informação. Segundo Araújo (2013) “tratar” a informação implica a execução de várias atividades, todas elas visando possibilitar ao usuário a recuperação da informação desejada com a maior acurácia, efetividade, precisão e rapidez possíveis. Estão implícitas nesta atividade, de acordo com Alonso (2012), as tarefas relacionadas à classificação, catalogação, indexação, dentre outras. Destas, a indexação se destaca sob o ponto de vista decisório, pois apresenta o desafio da análise de assunto e atribuição de termos que representem fielmente o documento objeto de sua investigação. A proposta deste trabalho é apresentar o resultado parcial da pesquisa de mestrado realizada com o objetivo de investigar os comportamentos informacionais envolvidos na atividade de análise de assunto executada na indexação. Essa preocupação baseou-se no aspecto relacionado à qualidade da indexação, prerrogativa que está condicionada à qualificação do indexador cabendo destacar a importância da imparcialidade no processo, apesar de se considerar ser quase inevitável que alguns elementos de julgamento subjetivo não afetem a realização da indexação. Utilizou-se como base teórica a Abordagem Clínica da Informação e metodologias de suporte que propiciaram uma análise profunda do fenômeno da informação pela via simbólico-afetiva, em especial o uso de imagens visando acessar a dimensão subjetiva dos



indivíduos utilizando o fenômeno da *poïesis* e o Teste Arquetípico de Nove Elementos – AT-9 que visa mapear o tipo de estrutura do imaginário por meio da qual o indivíduo se expressa. Os resultados apresentados referem-se a um recorte decorrente da análise simbólica efetuada sobre as imagens evocadas na entrevista e no AT-9, o que possibilitou compreender como os aspectos subjetivos podem se integrar às competências individuais e influenciar o processo decisório. Os resultados obtidos até agora com esse estudo apontam para a viabilidade de efetuar uma investigação do fenômeno da subjetividade humana em sua relação com a informação no contexto da cultura de urgência valendo-se de uma leitura interdisciplinar.

Palavras-Chave: Recuperação de informação; Análise de assunto; Subjetividade; Dimensões simbólico-afetivas

Referencias

ALONSO, Livia Ferreira Coutinho. A atividade de indexação: uma construção social da realidade. Dissertação (Mestrado). Escola de Ciência da Informação. Universidade Federal de Minas Gerais. Belo Horizonte, 2012

ARAUJO, Eliane Pawlowski de Oliveira. Tomada de decisão organizacional e subjetividade: análise das dimensões simbólico-afetivas no uso da informação em processos decisórios. Dissertação (Mestrado). Escola de Ciência da Informação. Universidade Federal de Minas Gerais. Belo Horizonte, 2013.

GARCIA, Joana Coeli Ribeiro. Recuperação da informação. DataGramZero - Revista de Ciência da Informação - v.8 n.6 dez/07. Disponível em http://www.dgz.org.br/dez07/F_I_com.htm. Acesso em 25/05/2014.



A NECESSIDADE DE DIVULGAÇÃO DOS OBJETIVOS E FUNCIONALIDADES DOS REPOSITÓRIOS INSTITUCIONAIS: RELATO DE EXPERIÊNCIA

Vilma Gravatá da Conceição
Bibliotecária da Universidade Federal da Bahia

Pôster: http://biredial.ufr.br/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/229/133

Resumo

Os Repositórios Institucionais (RIs) tem a finalidade de ofertar através de um único espaço virtual a produção acadêmica e científica das universidades, propiciando a visibilidade das mesmas, bem como, de seus pesquisadores, além de contribuir para a preservação da memória intelectual das instituições. Para atender as demandas informacionais dos usuários faz-se necessário, além da capacitação a divulgação das atividades realizadas nas bibliotecas e, também, os produtos que elas oferecem. Este estudo trata-se de um relato de experiência, desenvolvido após o treinamento sobre estratégias de busca de informação científica oferecido pelo setor de referência, da Biblioteca Universitária Reitor Macedo Costa, da Universidade Federal da Bahia, no qual observou-se que existe necessidade de uma divulgação mais estruturada com foco no Repositório Institucionais, pois os alunos apresentam interesse pelo assunto, porém com pouco conhecimento sobre os objetivos e funcionalidade do RIs. O problema da pesquisa é que ainda há uma carência de divulgação dos serviços do Repositório Institucional para otimização dos mesmos pela comunidade acadêmica das Universidades Brasileiras. A princípio foi realizado um levantamento bibliográfico sobre o processo de divulgação e disseminação dos objetivos e funcionalidades dos RIs. A pesquisa caracteriza-se como descritiva. A técnica adotada é da observação desenvolvida mediante a realização dos treinamentos de capacitação de usuários, a abordagem é qualitativa. Dessa forma, pode-se aferir que a educação dos usuários pode ser utilizada no sentido de divulgar os serviços da biblioteca incluindo os RIs abordando sua finalidade e função, vez que o processo de capacitação realizado por meios de treinamentos e cursos é capaz de promover a utilização efetiva dos produtos e serviços oferecidos pelas bibliotecas.

Palavras-Chave: Repositórios Institucionais, divulgação, visibilidade, capacitação de usuarios



Referencias

BLATTMANN, U; WEBER, C.. DSpace como repositório digital na organização. Revista ACB. Florianópolis, v.13, n. 2, p. 467-85, jul./dez, 2008.

LEITE, Fernando César L. Como gerenciar e ampliar a visibilidade da informação científica brasileira : repositórios institucionais de acesso aberto. Brasília : Ibict, 2009. 120 p.

SAYÃO, Luís. et. al (orgs). Implantação e gestão de repositórios institucionais : políticas, memória, livre acesso e preservação. Salvador : EDUFBA, 2009. Universidade Federal da Bahia. Repositório Institucional. Disponível em: <<https://repositorio.ufba.br/ri/>>. Acesso em: 27 maio 2014.

WEITZEL. S. R. O papel dos repositórios institucionais e temáticos na estrutura da produção científica. Em Questão, v. 12, n. 1, p. 51-71, 2006. Disponível em: <http://eprints.rclis.org/12101/1/weitzel_repositorios.pdf>. Acesso em: 29 maio 2014.



SEMBRANDO CONCIENCIA PARA COSECHAR CIENCIAS EN ACCESO ABIERTO- EXPERIENCIAS DE UNIVERSIDADES DEL NOROESTE ARGENTINO

Silvia Estela d'Empaire¹, Nicolas Federico Formoso²

¹Profesora Asociada, Departamento de Ciencias de la Computación, Facultad de Ciencias Exactas y Tecnología - Universidad Nacional de Tucumán; ²Universidad Nacional de Tucumán

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/231/107

Resumen

El incremento constante de la producción científica, el significativo crecimiento de la oferta académica a nivel de grado y posgrado en de las universidades argentinas en las últimas décadas, sumado el surgimiento de nuevos canales de comunicación a través de internet, han generado una urgente necesidad de crear un eficiente sistema que gestione la información por ellas producida, en todas sus formas través de Repositorios Digitales. Consiente de esta situación y en procura de satisfacer los requerimientos de las universidades públicas del país se lanzó desde el Ministerio de Ciencia y Tecnología e Innovación Productiva de la Nación Argentina en 2010 una convocatoria para desarrollo y mejoras de repositorios digitales institucionales de acceso abierto. Los proyectos dieron comienzo en 2012 y participaron 31 universidades del país agrupadas en regiones: Región Bonaerense, Metropolitana, Centro-Oeste, Centro-Este, Sur y el proyecto que nos ocupa: de la Región Noroeste con la participación de las Universidades Nacionales de Tucumán, Salta, Santiago del Estero y Jujuy. En este poster expondremos A) Los procesos de innovación-desarrollo llevados a cabo en el marco del proyecto de creación de repositorios digitales de acceso abierto, con la participación del equipo multidisciplinario formado por investigadores, informáticos y bibliotecarios de las cuatro universidades. Presentaremos también un extracto de las fases seguidas para resolver problemas que plantearon los objetivos específicos. Se contemplaron: Aspectos documentales y legales, organizacionales, técnicos, de difusión y capacitación. B) Los hitos significativos que reforzaron los resultados obtenidos.

Palavras-Clave: Acceso Abierto –Open Access- Repositorios Institucionales - DSpace - Universidades del Noroeste Argentino – Universidad Nacional de Tucumán



Citas

Barton, M; Waters, M. "Creating an Institutional Repository: LEADIRS Workbook" with sponsorship from The Cambridge-MIT Institute (CMI). Copyright © 2004 MIT Libraries

Abadal, Ernest (2012). "Acceso abierto a la ciencia." - Barcelona: Editorial UOC. (Colección El profesional de la información).

Suber, Peter – "Open Access" - MIT Press essential knowledge-2012 Massachusetts Institute of Technology

Dawson, Stephanie (2014) "Interview with Advisory Board member Peter Suber" <https://www.scienceopen.com/discussion/7ed7a938-a0bd-415c-b539-0b1d3ffbe4e8/4ff731f4-1e42-4b2f-8cf8-7865cab3394f>

Lynch, Clifford A. "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" 2003

Miguel, Sandra and Bongiovani, Paola et al. "Prospect for development of Open Access in Argentina". The Journal of Academic Librarianship, 2013, vol. 39, n. 1, pp. 1-2. [Journal Article (Print/Paginated)]

Watson, Kate – "SURVEY Institutional Repository: A Survey of Academic and Research Needs at The University of the Sunshine Coast" Aug 2006

Argentina. Honorable Cámara de Diputados de la Nación, *Proyecto de Ley 1927-D-2011 Creación de Repositorios Digitales Institucionales de Acceso Abierto, Propios o Compartidos*, Online. (2011). Available: <http://www1.hcdn.gov.ar/proyxml/expediente.asp?fundamentos=si&numexp=1927-D-2011>, (Junio 28, 2013).

De Giusti, Marisa Raquel – "Gestión de la información del conocimiento y del aprendizaje organizacional. Hacia la competitividad institucional" Available: http://sedici.unlp.edu.ar/bitstream/handle/10915/25157/Documento_completo.pdf?sequence=1 (Junio 28, 2013).



Abstract

The constant increase in scientific production, the significant growth of the academic offerings of undergraduate and graduate, in the Argentine universities in recent decades, added the rise of new channels of communication through the Internet, have created an urgent need to create an efficient system to manage the information produced by them, in all its forms through Digital Repositories. Aware of this situation and seeks to satisfy the requirements of public universities in the country, was launched by the Ministry of Science, Technology and Productive Innovation of Argentina in 2010 a call for development and improvement of institutional digital repositories of open access. The project began in 2012 and involved 31 universities grouped into regions: Buenos Aires Region, Metropolitan, Central West, Central East, South, and this, the project at hand: the Northwest Region with the participation of universities national de Tucumán, Salta, Santiago del Estero and Jujuy. This poster will discuss A) Innovation processes-development carried out under the project to create open access digital repositories, with the participation of a multidisciplinary team of researchers, computer scientists and librarians from the four universities. Also present an extract of the steps followed to solve problems raised by the specific objectives. Were contemplated: documentary and legal aspects, organizational, technical, information and training issues. B) significant milestones which strengthened the results obtained

Keywords: Open-Access- Institutional Repositories - DSpace - Universities of Northwest Argentina - National University of Tucumán



PRESERVACIÓN DE LA MEMORIA DEL SENA

Martha Luz Gutiérrez Ortega, Adriana Rincón Avendaño
Servicio Nacional de Aprendizaje SENA

Pôster: http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/viewFile/234/114

Resumen

El Servicio Nacional de Aprendizaje SENA fundado en 1957, está encargado de cumplir la función que le corresponde al Estado, de invertir en el desarrollo social y técnico de los trabajadores colombianos, ofreciendo y ejecutando la formación profesional integral, para la incorporación y el desarrollo de las personas en actividades productivas que contribuyan al desarrollo social, económico y tecnológico del país. (Ley 119 de 1994). Desde los años 60, el SENA ha conformado una colección bibliográfica y audiovisual variada y valiosa como parte de las acciones involucradas en el proceso de Formación Profesional Integral, que dan cuenta de sus experiencias pedagógicas didácticas y buenas prácticas en materia de formación para el trabajo. A partir del 2006 el Sistema de Bibliotecas SENA, formuló el Proyecto de preservación de la memoria institucional, por medio del cual se ha identificado y reconstruido, los diferentes modos de formación Aprendizaje, que ha adoptado la entidad, en su propósito de actualizar tecnológicamente a los trabajadores activos, denominada "formación complementaria" y la formación de Técnicos y de Tecnólogos. Originalmente los materiales de apoyo didáctico estaban conformados por "Series metódicas y colecciones básicas ocupacionales" elaboradas con el apoyo de CINTERFOR. Posteriormente como opción educativa, se realizaron programas radiales dirigidos a la población Urbana y Rural y como producto del Centro Experimental de Formación Abierta y a Distancia (CEFAD), se editaron cartillas de autoformación y audiovisuales de apoyo a la formación. Igualmente como producto de los diferentes convenios nacionales e internacional se encuentran estudios en las diferentes áreas de formación, como es el caso de los estudios de cosecha y pos cosecha que se editaron entre 1995 y el 2003 en convenio con el Reino Unido. El propósito de la presente ponencia es mostrar cómo se recupera, preserva y divulga la producción intelectual del SENA. Divulgar la producción histórica y actual de la Institución que se constituye en la Memoria y el patrimonio de la Nación un aporte al mundo. **Objetivos Específicos:** Recuperar y seleccionar el material bibliográfico y audiovisual que reposa en los acervos de las bibliotecas físicas para consolidar el know how y patrimonio de la entidad. Organizar y digitalizar el material bibliográfico y audiovisual en formatos estándar para garantizar la preservación, de acuerdo a los lineamientos nacionales e internacionales. Publicar y dar acceso abierto a la producción intelectual del SENA en el Repositorio Institucional. La Colección SENA se ha presentado en diferentes escenarios como, en la Feria Internacional del Libro en Bogotá, (2010, 2011 y 2012), en el Valle y Medellín (2011). Presentación en Bibliotic, en el Encuentro nacional de Bibliotecas (2011). Presentación de ponencia en el Seminario Internacional de Conservación de Archivos Audiovisuales organizado por Fonoteca Nacional en



Bogotá año 2012. En el 2013, la entidad toma la decisión de implementar y personalizar la herramienta tecnológica de código abierto Dspace, para el registro, actualización, organización, administración, búsqueda y recuperación de los objetos digitales del SENA. A la fecha, se han digitalizado 8.000 archivos de los cuales se han catalogado e ingresado 1.446 archivos para su visibilización en el Repositorio Institucional <http://repositorio.sena.edu.co/>.

Citas

Memoria del Mundo – directrices para la salvaguardia del patrimonio documental. UNESCO (edición revisada 2002).

SENA 50 años de formación profesional.

Convención Nacional de Centros Binacionales – 11 y 12 de Octubre de 2007. Trujillo. Perú - La Importancia de la Digitalización de Archivos para la Biblioteca http://eprints.rclis.org/11912/1/La_importancia_de_la_digitalización_de_archivos_para_la_bi....pdf.



Índice de Autores

Nomes	Página
A	
Adriana Rincón Avendaño	399
Alexa Ramírez Vega	353
Alline de Sousa	383
Aline dos Santos Jacob	24, 256, 365
Aline Silva	369
Ana Carla Cunha Nascimento	363
Ana Maria Mielniczuk de Moura	389
Angélica Conceição Dias Miranda	381
Antonio Felipe Razo Rodríguez	9
Antonio Rubio-Montero	166
Ariadne Chloe Mary Furnival	272, 367
Ariel Jorge Lira	58, 81, 201
B	
Bianca Miranda de Almeida	272
Bruce Becker	166
Bruno Marçal	349
C	
Carla Carrubba	166
Carla Lopes Ferreira	38
Christos Kanellopoulos	166
Cícera Henrique da Silva	369
Clayton Martins Pereira	387
Claudia A. Tocantins	153
Cláudia Oliveira de Moura Bueno	38
Claudio Paixão Anastácio de Paula	392



Nomes	Página
Concha Mosquera-de-Arancibi	94
Cristiane Castro Soares	381
D	
Daniel Flores	285, 373
Daniela Gralha de Caneda Queiroz	389
Daniela Maciel Pinto	359
Danilo Formenton	215
Denise Nacif Pimenta	369
Denyson Mota	357
Diane Cátia Tomasi	376
Dhion Carlos Hedlund	285, 373
Dora Isabel Diaz Susa	342
E	
Edison José Correa	379
Eliane Pawlowskii Oliveira Araújo	379, 392
Elis Gabriela Copa dos Santos	367
Emilio Lorenzo-Gil	49
Ester Aparecida Lima de Souza	296
Evandro Anacleto Abreu Ferreira	373
F	
Facundo Gabriel Adorno	58
Federico Ruggieri	166
Fernanda de Souza Monteiro	256, 365
Fernanda Weschenfelder	385
Franco Agustin Terruzzi	201
G	
Gabriela Silva da Rosa	381
Giuseppina Inserra	166
Gonzalo Luján Villarreal	81, 201
Gonzalo Rubén Alvarez	350



Nomes	Página
H	
Helen Beatriz Frota Rozados	350
Henrique Machado dos Santos	285, 373
I	
Isadora Turcatel	361
J	
Jackson da Silva Medeiros	391
Jacqueline Pawlowski Oliveira	379
Jean Gerald Francis Banon	387
Johan Espinoza Rojas	328
Jorge Alberto Soares Cruz	373
Jorge Polanco-Cortés	328
Jorge Santa Anna	122
José Renato Margarido Galvão	383
Jose Texier	81
Juliana Guerra de Barcellos	350
K	
Kátia V. Cavalcante	153
Kostas Koumantaros	166
L	
Laís Pereira de Oliveira	382
Laura Vilela Rodrigues Rezende	244, 382
Linamaría Pinzón Valencia	371
Luciana de Souza Gracioso	215
Luis Martínez-Uribe	104
Luis A. Núñez	166
Luis Guilherme Gomes de Macena	369
M	
Manuel Rodriguez Pascual	166
Marciana Leite Ribeiro	387



Nomes	Página
Márcia Medeiros de Carvalho Mendo	383
Márcio Bezerra da Silva	363
Maria Cristina Monteiro Tasca	383
Maria da Conceição Carvalho	369
Maria Helena Machado de Moraes	381
Maria João Amante	310, 349
Maria Rizoneide Negreiros Araújo	379
Maria Silvério da Silve Siqueira	38
Marisa Raquel De Giusti	58, 81, 201
Martha Luz Gutiérrez Ortega	399
Mayara Campos Souza	363
Meilyn Garro-Acón	328
Morena Pereira Porto	141
N	
Nair Yumiko Kobashi	357
Nicolas Federico Formoso	396
O	
Ognjen Prnjat	166
Onivaldo Rosa Junior	24
P	
Paola Carolina Bongiovani	104
Paula Balbis Garcia	141
Paula Porto Gauterio	381
Perla Olivia Rodríguez Reséndiz	185
R	
Rafael Mayo-García	166
Rafael Porto da Rocha	361
Rejane Machado	369
Rita Ricceri	166
Roberto Barbera	166



Nomes	Página
Rosalina Vázquez Tapia	9
Rosemary Mendonça Martins Fernandes	383
Rubens da Costa Silva Filho	376
S	
Salma Jalife Villalón	166
Samile Andréa de Souza Vanz	350, 355
Sergio Nieto-Caramés	49, 94
Silvia Estela d'Empaire	396
Sonia Cruz Riascos	244, 382
Sônia Elisa Caregnato	391
Sonia Maria Valencia Grajales	342
Soraya Alves Lacerda	256, 365
Suelen Spíndola Bilhar	355
T	
Teresa Maria Segurado	310, 349
V	
Valéria Gameleira da Mota	385
Vera Regina de Aquino Vieira	141
Vilma Gravatá da Conceição	394
Vinícius de Araújo Oliveira	256
Viviane Santos de Oliveira Veiga	369