



20 al 30 de abril de 2017

## DIRECTORIO INTEGRADO DE MATERIALES EN ACCESO ABIERTO (DIMA)

## INTEGRATED DIRECTORY OF MATERIALS IN OPEN ACCESS (DIMA)

*Experiencias y recursos en educación virtual 2.0. Los cursos MOOC abiertos masivos en línea: Comunicación de experiencias, evaluación e impacto de esta nueva tendencia*

*Yanedi Abreu Bartomeo<sup>1</sup>, Francisco Lee Tenorio<sup>2</sup>, Luis Dominguez Cruz<sup>3</sup>, Raúl Gonzalo Torricella Morales<sup>4</sup>*

<sup>1</sup> Universidad de las Ciencias Informáticas (UCI), Cuba, [yabreub@uci.cu](mailto:yabreub@uci.cu)

<sup>2</sup> Ministerio de Educación Superior (MES), Cuba, [fleet@reduniv.edu.cu](mailto:fleet@reduniv.edu.cu)

<sup>3</sup> Dofleini Software, Cuba, [ldominguezc@gmail.com](mailto:ldominguezc@gmail.com)

<sup>4</sup> Ministerio de Educación Superior (MES), Cuba, [torri@reduniv.edu.cu](mailto:torri@reduniv.edu.cu)

### RESUMEN

El presente trabajo recopila los antecedentes, características, desarrollo, impacto y expectativas del directorio integrado de materiales en acceso abierto (DIMA). DIMA parte de la necesidad de facilitar la recuperación de información científica a estudiantes, profesores e investigadores y para ello se basa en asumir el modelo de acceso abierto como principio de las funcionalidades desarrolladas. El objetivo del



20 al 30 de abril de 2017

presente trabajo es describir los beneficios y posibles impactos de la aplicación DIMA mediante el análisis de sus funcionalidades y arquitectura, para evaluar su aporte al impulso y aprovechamiento del modelo de acceso abierto en un ambiente orientado a la web social y su contribución económica y social en la red de universidades del MES. La red universitaria del Ministerio de Educación Superior (MES) en Cuba es el entorno donde se han realizado pruebas a la aplicación. El principal aporte de DIMA está vinculado a la posibilidad de adquirir de forma libre y gratuita cerca del 30% de la información publicada en fuentes que requieren altos costos de suscripción como es el caso de ISI Web del Conocimiento (*Web of Knowledge*). También cabe señalar el impulso al modelo de acceso abierto y al crecimiento científico.

#### **Abstract**

The present work compiles the antecedents, characteristics, development, impact and expectations of the integrated directory of materials in open access (DIMA). DIMA starts from the need to facilitate the retrieval of scientific information to students, professors and researchers and for this is based on assuming the open access model as a principle of the functionalities developed. The objective of the present paper is to describe the advantages and the impact of the DIMA application through the analysis of its functionalities and architecture, to make an evaluation about the open access model approach in a social web environment and the economical contribution to the universities network of the Superior Ministry of Education. The university network of the Ministry of Higher Education (MES) in Cuba is the environment where the application has been tested. The main contribution of DIMA is linked to the possibility of acquiring, free of charge, about 30% of the information published in sources that require high subscription costs, such as ISI Web of Knowledge. The momentum for the open access model and for scientific growth should also be noted.

#### **Palabras Clave**



20 al 30 de abril de 2017

Acceso Abierto; Biblioteca digital, Dublin Core; OAI-PMH; Recuperación de información

### Key words

Open Access; Library digital, Dublin Core; OAI-PMH; Information Retrieval

## INTRODUCCIÓN

El modelo de acceso abierto (OA, por sus siglas en inglés) a la información académica y científica a lo largo del tiempo ha experimentado un crecimiento tanto cuantitativo como cualitativo. Los principales países y productores de información científica se han sumado de forma sustantiva a este movimiento. Uno de los catalizadores fue precisamente el incremento de los precios de las revistas científicas Martín, Millán, & Villavicencio, 2009 y la exigencia de los propios investigadores en aras de que sus trabajos fuesen accesibles y que su impacto se reflejase en la sociedad Melero & Abad, 2008. El caso más reciente que da un espaldarazo al acceso abierto (OA) viene del premio nobel de medicina en el 2013, Randy Schekman, quien plantea el dilema que hoy en día enfrentan las principales revistas científicas y la posibilidad que ofrece el “amplio abanico” de revistas en acceso abierto.

Económicamente asumir el OA presenta muchas ventajas, asociadas precisamente a los altos precios que hoy implica el acceso a la información de frontera. Actualmente cerca del 30% (media entre las diversas áreas del conocimiento) de los materiales que se encuentra en los grandes consorcios y editoriales de información científica, se pueden encontrar en alguna modalidad de acceso abierto Gargouri, Lariviere, Gingras, Carr, & Harnad, 2012. Otra ventaja está en el hecho de que las instituciones que asumen el OA por mandato logran un crecimiento científico de hasta el 60% Gargouri et al., 2010.

En el caso de América Latina también se hace patente el fortalecimiento del modelo [6]. Este hecho contrasta con el caso de Cuba, donde solo el sector de la salud ha logrado incorporarse de forma efectiva M. M. Fernández, 2012. Las principales causas que se identifican del pobre aprovechamiento del OA en Cuba están relacionadas con el limitado conocimiento del modelo y la poca consciencia de sus beneficios.

El asumir el modelo OA no es solo exponer revistas u otro tipo de publicaciones bajo licencias que permitan su distribución gratuita y libre para reproducir, citar y utilizar un recurso. Se debe contar con un soporte tecnológico que permita el autoarchivo y la diseminación del conocimiento. Se necesita implementar el protocolo OAI-PMH (Open Archive Initiative – Protocol for Metadata Harvesting) para garantizar la interoperabilidad independientemente de la diversidad tecnológica que pueda existir,



20 al 30 de abril de 2017

se requiere de proveedores de datos y proveedores de servicios, los componentes fundamentales del protocolo Bartomeo, Cruz, & Fernández, 2011.

El empleo de proveedores de servicios OAI-PMH en una institución facilita la recuperación de información científica, porque permite recolectar los documentos procedentes de proveedores de datos (revistas científicas o repositorios institucionales) y concentrarlos, permitiendo que estos sean recuperados desde una interface de usuario única. Para países, instituciones o comunidades, donde el acceso a internet presente limitantes el proveedor de servicios adquiere mayor significado.

Los sistemas informáticos más empleados para proveedores de datos y de servicios son OJS (Open Journal System) y Dspace para revistas y repositorios respectivamente M. M. F. Fernández, Cruz, & Bartomeo, 2013. Estos sistemas son muy buenos para la gestión de información, pero no incluyen prestaciones para el trabajo en comunidades, elemento primordial en la web actual, que cada vez es más social.

El Directorio Integrado de Materiales en Acceso Abierto (DIMA) es una aplicación web basada en la arquitectura del sistema de gestión de contenidos Drupal. Surge para suplir la necesidad de facilitar el acceso a la información científica. Tiene el objetivo de impulsar y aprovechar el modelo OA. Es una aplicación genérica, pero en este caso específicamente se ha desplegado para su uso en la Red de Universidades del Ministerio de Educación Superior (MES) en Cuba.

El objetivo del presente trabajo es describir los beneficios y posibles impactos de la aplicación DIMA mediante el análisis de sus funcionalidades y arquitectura, para evaluar su aporte al impulso y aprovechamiento del modelo de acceso abierto en un ambiente orientado a la web social y su contribución económica y social en la red de universidades del MES.

## **DESARROLLO**

Debe quedar claro, que el OA, no implica que los documentos sean gratis y libres, debido a que muchas veces se hace necesario pagar licencias, comprar los artículos de algunas revistas, pues muchos documentos se encuentran bajo los derechos de las editoriales. El OA es un cambio de paradigma, una nueva forma de divulgar la información científica.

El contexto general en el que se integra este movimiento es el denominado "conocimiento libre", que incluye también el software libre, o la cultura libre en general (wikipedia, música, etc.), siendo el OA a la ciencia la parte que se ocupa de los contenidos científicos. Es un movimiento que reclama la construcción de un dominio público para la ciencia y la cultura, que permita la difusión y reutilización del conocimiento y, por extensión, un rápido progreso científico y cultural Abadal, 2012.



20 al 30 de abril de 2017

Una de las alternativas que utilizan varios buscadores para recuperar la información de contenido científico-técnico son los repositorios y las revistas de acceso abierto (OA del Inglés Open Access). Los repositorios OA se basan, fundamentalmente, en poner a disposición de toda la comunidad científica y el público general, los artículos científicos y otros materiales docentes e investigativos a través de su publicación en revistas de acceso abierto y el depósito de estos en repositorios institucionales o temáticos de acceso abierto M. M. F. Fernández, Cruz, & Romero, 2012 Tarragó, 2007.

Las universidades, dentro de este movimiento, se convierten en uno de los proveedores de información más importante por la cantidad y calidad de documentación que generan. Actualmente muchas universidades, incluyendo aquellas consideradas como las más prestigiosas, exponen sus contenidos en repositorios institucionales con características OA. Algunas universidades y centros de estudios e investigación de América Latina, se unen a estas iniciativas de exponer sus publicaciones en repositorios de OA, además de crear sus propios repositorios. A causa, del alto costo de las revistas y publicaciones de corte científico. Por ello, es de gran importancia, que Cuba, o cualquier país, cuente con un directorio integrado de materiales en acceso abierto, como lo es el DIMA; para apoyar el proceso científico-educativo en las universidades.

### ***Características y beneficios del DIMA***

La subscripción a los grandes sistemas de gestión de información científica es en extremo costosa (en el orden de los millones). DIMA permite adquirir parte de esa información o investigaciones paralelas sin costo alguno, basándose en la premisa de que muchos autores publican sus trabajos también en alguna modalidad de acceso abierto (repositorio institucional o revista científica).

En la actualidad es práctica en muchas áreas del conocimiento publicar en fuentes libres los preprints de trabajos que finalmente terminan siendo publicados en las más importantes revistas científicas. Los casos más ilustrativos son: Arxiv en investigaciones relacionadas con las matemáticas, la física y la ciencia de la computación; RePec en temas de economía, BioMed Central en el área de la medicina y la biología y Rclist en ciencias de la información.

El crecimiento científico de las instituciones es otro de los beneficios, tanto económico como social. DIMA no implica que una institución que haga uso de la plataforma tenga un mandato para asumir el acceso abierto, pero si contribuye a que se citen trabajos que están en ese modelo. Esto aumenta la posibilidad de adquirir citas y contribuye al crecimiento científico.

El acceso a internet es otra área donde se incide directamente. Con la aplicación se logra concentrar la información que sea de interés y de esta forma una institución dedicada a la investigación puede disminuir considerablemente la necesidad de acceder a internet para la búsqueda de información, fomentando el uso de una red



20 al 30 de abril de 2017

local. Esta alternativa sería óptima para entes donde el acceso a internet supone un problema económico o de infraestructura.

Se puede decir, que contribuye de forma libre al acceso de la información científica. Esto repercute en mayor calidad de las investigaciones y mayor socialización de los resultados, algo de lo que adolece la ciencia actual, independientemente del desarrollo existente en el sector de la información y las comunicaciones.

### ***Funcionalidades del DIMA***

Entre las principales funcionalidades del DIMA, se encuentra la recuperación de acervos. Los mismos se indexan de diferentes fuentes de acceso abierto, incluyendo SCOPUS, EBSCO, DOAJ, entre otras. Además tiene integrado un módulo que normaliza el metadato autor, posibilitando que los nombres de autores no se encuentren repetidos y escritos de forma única. Los estudiantes, profesores, podrán acceder a documentos de alto valor científico mediante el DIMA.

Recuperación de información: Los tipos de contenidos fuente-colección-documento están relacionados y tributan a la recuperación de información, tanto Adhoc como por navegación. En el caso de la recuperación Adhoc, se cuenta con un buscador que previamente requiere de la indexación de los contenidos recolectados en un índice invertido. Para la recuperación por navegación se inicia a partir del enlace fuentes, ubicado en la parte superior del sitio.

Eliminación de duplicados: Se realiza como una tarea automática, utilizando el hook\_cron de la arquitectura de Drupal. Esta funcionalidad busca contenido duplicado por fuentes, analiza una fuente cada una hora. El control de duplicados se realiza por comparación de similitud entre documentos, utilizando el modelo vectorial para la recuperación de información.

Presentación: La presentación del sistema incluye un diseño adaptable (responsive design) que permite una correcta visualización en diferentes dispositivos (computadoras de escritorio, laptops, tablets y smartphone). Se hace uso del framework Bootstrap para la incorporar estilos a tablas, menús, formularios y botones.

Exportar y compartir información: Los metadatos de los documentos pueden ser exportados a los principales manejadores de referencia (Endnote, Zotero), además el recurso puede compartirse en las principales redes sociales.

Recolección de documentos: El harvester recolecta solamente los metadatos de los documentos, por otra parte los administradores del sistema tienen la posibilidad de configurar la recolección de los documentos físicos (doc, pdf, ppt). En entornos sin limitaciones para el acceso a internet no es aconsejable utilizar esta opción. En el caso de que existan limitaciones para el acceso a internet puede tomarse esta alternativa para facilitar el acceso al texto completo desde la red local. Tiene la limitante del gran espacio de almacenamiento que requiere. Esta funcionalidad



20 al 30 de abril de 2017

también es de alto costo computacional por lo que se realiza por lotes, mediante tareas cron.

## **CONCLUSIONES**

La principal conclusión que arroja el trabajo es que se hace necesario el impulso del modelo de acceso abierto, sobre todo en países con limitaciones de acceso a información científica. Los proveedores de servicios OAI-PMH son una solución tecnológica que favorece ese objetivo. DIMA como solución informática cuenta con una arquitectura muy flexible que permite que el sistema pueda ser adaptado y generalizado con facilidad en diversos escenarios. La recolección, indexación y los métodos de recuperación de información que posee DIMA permiten aprovechar recursos en acceso abierto que de otra forma requerirían un esfuerzo considerable por parte de los usuarios, para encontrar la información, sobre todo aquellos que se inician en la investigación. La concentración de metadatos presente en DIMA contribuye a disminuir la total dependencia de internet para realizar una investigación. Esto adquiere mayor relevancia cuando se trata de escenarios con limitaciones en la infraestructura de comunicaciones.

## **REFERENCIAS**

<http://www.ub.edu/bid/20meler2.htm>



20 al 30 de abril de 2017

Yanedi Abreu Bartomeo: Graduado de Ingeniería en Ciencias Informáticas en el año 2007. Es profesor asistente. Ha impartido clases en pregrado de: Ingeniería de Software, Programación Web, Introducción a las Ciencias Informáticas, Factibilidad de proyectos informáticos y Comercio electrónico. En postgrado ha impartido el curso de Recuperación de Información. Se encuentra matriculada en la maestría de informática aplicada. Tiene varias publicaciones sobre acceso abierto, biblioteca digital y normalización del nombre de autor en repositorios científico. Ha cursado varios cursos de postgrado. Tiene conocimientos de tecnologías web tales como el CMS Drupal, HTML/CSS, JS, Query, PHP, WordPress, Prestashop, entre otros.

#### Foto de los autores



Ing. Yanedi Abreu Bartomeo



Drc. Francisco Lee Tenorio



Ing. Luis Dominguez Cruz



Drc. Raúl Torricella Morales