

Limitaciones en la distribución de contenidos entre repositorios de acceso abierto en América Latina y el Caribe

Limitations in the distribution of contents among Latin American and Caribbean open access repositories

Maikel Manuel Fernández Fernández, Luis Domínguez Cruz, Yanedi Abreu Bartomeo

Universidad de las Ciencias Informáticas. La Habana, Cuba.

RESUMEN

El presente trabajo muestra un estudio de 58 fuentes de acceso abierto correspondientes a 14 países de América Latina y el Caribe. Se busca identificar algunos problemas relacionados con la estandarización del protocolo OAI-PMH en los repositorios de acceso abierto en América Latina y el Caribe y cómo estos afectan la distribución de contenidos. El objetivo que el trabajo persigue es identificar las limitantes que afectan el intercambio de contenidos entre fuentes de acceso abierto en América Latina y el Caribe mediante el estudio de repositorios institucionales y de revistas de corte científico que hacen función de proveedores de datos. Las fuentes seleccionadas responden a revistas de corte científico y a repositorios de universidades. El estudio fue realizado con el empleo del software OHS (*Open Harvester System*) y tiene en cuenta aspectos como la tecnología empleada, los estándares de metadatos y la forma de representación de la información, además del uso de los verbos que establece el protocolo OAI-PMH. Se hizo una observación del funcionamiento de algunos repositorios de acceso abierto y se probaron las respuestas de estos ante las peticiones de los verbos. Los resultados obtenidos muestran diversidad, tanto tecnológica como en la forma de aplicar los estándares, y reflejan algunas de las problemáticas a las que se enfrenta el movimiento de acceso abierto en el área de estudio y cómo esto afecta la interoperabilidad necesaria para automatizar el intercambio de contenidos.

Palabras clave: repositorios; protocolos de comunicación; estándares de metadatos; OAI-PMH; proveedores de datos; acceso abierto.

ABSTRACT

A review was conducted of 58 open access sources from 14 Latin American and Caribbean countries, with the purpose of identifying problems related to the standardization of the OAI-PMH protocol in Latin American and Caribbean open access repositories, as well as the way in which these problems affect the distribution of contents. The study was aimed at identifying the limitations affecting the exchange of contents between Latin American and Caribbean open access sources through the study of institutional repositories and scientific journals which perform the function of data providers. The sources selected were scientific journals and university repositories. The study was based on the software OHS (Open Harvester System) and considered aspects such as the technology used, metadata standards and the manner in which information was represented, besides the verbs established by the OAI-PMH protocol. The functioning of some open access repositories was observed and tests were conducted of their response to the requests of verbs. Results show diversity both in the technology and in the way standards are applied, as well as some of the problems faced by the open access movement in the study area and their impact on the interoperability required to automate the exchange of contents.

Key words: Repositories; communications protocols; metadata standards; OAI-PMH, data providers; open access.

INTRODUCCIÓN

El intercambio de acervos entre fuentes de información es algo común en la actualidad. La informática y la ciencia de la información han desarrollado estándares y tecnologías para facilitar este propósito. El movimiento de acceso abierto (OA del Inglés Open Access), apoyado por el incremento de las publicaciones OA y el aumento de su visibilidad, ha cobrado fuerza y hoy es una alternativa viable para muchos países e instituciones.^{1,2}

En el año 2002 la declaración de Budapest y en el 2003 las de Bethesda y Berlín crearon las primeras pautas; La primera definió el concepto de OA y planteó la necesidad de los repositorios institucionales y las revistas de acceso abierto como estructuras para alcanzar el OA. Bethesda y Berlín afirman la declaración anterior y aportan el compromiso de avalar el modelo y desarrollar las herramientas necesarias para su implantación.^{3,4} La diversidad en tecnología y en la forma de describir y distribuir los documentos en un principio fue un verdadero freno. Con el tiempo se desarrollaron estándares y protocolos para normar los procesos de exponer y recibir los contenidos a través de internet. El principal resultado para propiciar el intercambio de conocimiento es el protocolo OAI-PMH. Sus siglas responden a: *Open Archives Initiative Protocol for Metadata Harvesting*^a. Los sistemas que intervienen en el acceso a los metadatos de distintas fuentes son divididos en proveedores de datos (DP del Inglés *Data Providers*) y proveedores de servicios (SP del Inglés *Service Providers*).

Los DP son software que permiten describir los documentos y que responden a peticiones de los SP. Mientras que los SP analizan las respuestas de los DP y gestionan los resultados, son empleados para indexar contenidos de diversos DP para luego ofrecer servicios de valor añadido sobre los metadatos recolectados. Permiten que los productores de información puedan hacer un intercambio de información, que aumentan los límites en el acceso a un documento. OAI-PMH es una tecnología simple que realiza la comunicación mediante el intercambio de documentos XML^b vía HTTP^c. Además aboga por el uso del Dublin Core como estándar de metadatos para describir los documentos en las fuentes.⁵⁻⁷ Ha sido determinante en el éxito del OA, principalmente porque sus pautas han logrado simplificar el problema del intercambio de datos libremente, y han dado paso a que diversos sistemas informáticos implementen el intercambio de información independientemente de la tecnología empleada.

Pareciera que este protocolo resuelve todo el problema, pero lamentablemente no es así; aún quedan muchos aspectos donde el factor humano es determinante. Si bien los DP han evolucionado rápidamente y hoy se cuenta con varias aplicaciones para este fin, no se puede decir lo mismo de los SP. La principal causa es la falta de estandarización en la estructura de los XML que son intercambiados entre los repositorios, que persiste a pesar de las normas ya establecidas. La cantidad de repositorios y revistas con un modelo OA ha aumentado considerablemente. América Latina y el Caribe también se han sumado a este movimiento, en el que las universidades determinantes por su alto nivel de producción de información son de corte académico y científico. Países como Brasil, Argentina y México marchan a la vanguardia,^{8,9} también proyectos como ScIELO^d y RedAlyC^e en la recuperación de datos de revistas científicas.¹⁰ Este auge es muy halagüeño, pero se debe prestar especial atención a las diferencias existentes que pueden atentar contra la interoperabilidad y dar al traste con el aprovechamiento de estos fondos documentales.

Aspectos como el manejo de las palabras clave, el uso de las variables que propone OAI-PMH, el uso que se haga de la tecnología y de los estándares de metadatos son determinantes y son analizados en el presente trabajo para identificar las diferencias existentes y sacar conclusiones que permitan afrontar el desarrollo de proveedores de servicios. El estudio incluye 58 DP donde se representan 14 países de América Latina y el Caribe. Las fuentes seleccionadas están todas registradas en ROAR^f (<http://roar.eprint.org>) y muestran altos niveles de actualización hasta el primer trimestre de 2012.

El objetivo que el trabajo persigue es identificar las limitantes que afectan el intercambio de contenidos entre fuentes de acceso abierto en América Latina y el Caribe mediante el estudio de repositorios institucionales y revistas de corte científico que hacen función de proveedores de datos.

PROVEEDOR DE DATOS

El proveedor de datos es una de las estructuras que propone el protocolo OAI. Es la encargada de crear, conservar y depositar los recursos en repositorios, además de extraer sus metadatos para exponerlos. Para garantizar la interoperabilidad, los proveedores de datos describen su información usando Dublin Core como estándar de metadatos,^{11,12} lo cual no quiere decir que no se pueda complementar con el uso de otros estándares.

La estandarización no solo se logra con la descripción homogénea de los contenidos, sino también fijando determinados patrones para la comunicación. Los

proveedores de datos OAI responden a seis verbos ya establecidos y emiten una respuesta en formato XML para cada uno de estos.¹³⁻¹⁵ Ellos son:

- *Identify*: pregunta por la identidad de una fuente; se utiliza frecuentemente para identificar si la fuente es un proveedor de datos y si está activo; también se obtienen otras informaciones como el formato en que se manejan las fechas.
- *ListMetadataFormats*: pregunta por los estándares de metadatos con que están descritos los documentos en la fuente encuestada.
- *ListSet*: solicitud sobre los *Sets*, o conjuntos en los que está agrupada la información.
- *ListIdentifier*: pide un recurso del que se conoce su identificador.
- *ListRecords*: pide el listado de los documentos.
- *Getrecord*: pide un recurso específico.

Estos verbos se complementan con otras variables como *from* y *until* para definir rangos de fechas en las peticiones *ListRecords*, *resumptionToken* como señal de reanudación para respuestas muy extensas y *set* para filtrar por colecciones o conjuntos, también en peticiones *ListRecords*. Para efectuar esta tarea de responder a las diversas peticiones el proveedor de datos debe contar con un generador de XML, un sistema de validaciones, un servidor Web y un servidor de base de datos, además de una interfaz para extraer los metadatos. Es vital que los proveedores de datos respondan de manera análoga para que los proveedores de servicios puedan funcionar de forma eficiente y para que se pueda aprovechar toda la información expuesta de forma libre, la cual puede ser accesible, pero las limitantes tecnológicas la hacen inaccesibles.

PROVEEDOR DE SERVICIOS

La función de este componente es aprovechar al máximo los metadatos que aportan los proveedores de datos, basado siempre en el enfoque *Harvester*^a que propone el protocolo OAI-PMH.^{16,17} Un proveedor de servicios tiene muchos usos derivados del indexado de información^b, entre ellos, el de implementar sistemas de recuperación de información.^{18,19} El desarrollo de un proveedor de servicios incluye:

- La selección y validación de proveedores de datos.
- El control de flujo.
- La planificación de recolecciones.
- La recolección de metadatos.
- La detección de contenido duplicado.
- El análisis de la granularidad de las fechas.
- Interfaz de usuario para la interacción con el público.

Estas herramientas están pensadas para la interacción con los usuarios, pero su éxito depende en gran manera de la efectividad de los proveedores de datos.

MÉTODOS

Los 58 repositorios seleccionados para el estudio están registrados en ROAR y son actualizados dinámicamente. Se trató de representar a la mayoría de los países de la región y además de que existiese diversidad tecnológica y de estándares de metadatos, independientemente de que todas las fuentes hagan uso del estándar Dublin Core (oai_dc¹) por exigencia del protocolo. En la [figura 1](#), se muestra la cantidad de repositorios por países que integran la muestra.

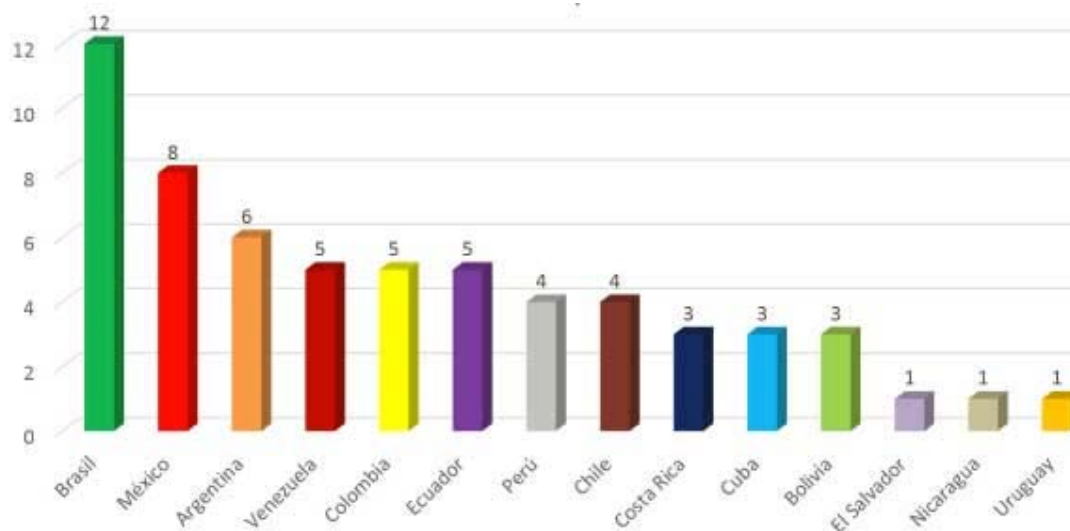


Fig. 1. Distribución por países de la muestra de proveedores de datos tomada para el estudio.

En el presente trabajo se hace uso de la herramienta Open Harvester Systems (OHS), sistema desarrollado por Public Knowledge Project (PKP) y que se encarga del indexado de metadatos,²⁰ haciendo función de proveedor de servicios. Esto permitió probar la compatibilidad de los proveedores de datos e identificar las principales limitantes. Se midieron cinco aspectos de cada una de las 58 fuentes, los cuales presentan la mayor probabilidad de errores de estandarización a la hora de ser cosechados los registros entre un repositorio y un proveedor de servicios:

- *Granularidad de la fecha:* el formato en que se encuentra la fecha en los documentos. Es importante que los documentos cosechados por un proveedor de servicios desde diferentes fuentes tengan el mismo formato de fecha, para facilitar al usuario las búsquedas. De encontrarse diferentes formatos será necesario realizar tareas de estandarización.
- *Formato de metadatos soportados:* pueden ser oai_dc, rdf, ore, mets. Mientras más formatos soporte el proveedor de datos más alta es la posibilidad de que sea indexado desde diferentes proveedores de servicios y a su vez más difundidos son los documentos.
- *Presencia de colecciones:* la existencia de colecciones en los proveedores de datos permite a los proveedores de servicios realizar la cosecha selectiva de documentos.
- *Token:* este término se refiere al elemento $\langle resumptionToken \rangle^1$, que debe estar presente al final de cada solicitud de indexación de documentos. Así el proveedor de servicios puede saber cuál es la siguiente página de documentos que debe

pedir al proveedor de datos. Sin este *token* es imposible recibir una colección completa.

- *Palabras clave*: el formato en que se encuentran las palabras clave de los documentos almacenados en el proveedor de datos.

Mediante la utilización del OHS en función de proveedor de servicios se encuestaron cada una de las fuentes para obtener la información estadística e identificar los principales problemas. La lógica del funcionamiento en cada caso es la siguiente:

- *Granularidad de la fecha*: se consulta la URL base del proveedor de datos o repositorio concatenándole *?verb=Identify*. Luego debe observarse la llave *granularity* en el XML resultante (ejemplo: <http://tesis.repo.sld.cu/cgi/oai2?verb=Identify>).
- *Formato de metadatos soportados*: se consulta la URL base del repositorio concatenándole *?verb=ListMetadataFormats*. Aquí debe prestarse atención a los valores que aparecen en las llaves *metadataPrefix* del xml resultante (ejemplo: <http://tesis.repo.sld.cu/cgi/oai2?verb=ListMetadataFormats>).
- *Presencia de colecciones*: para conocer si existen colecciones se concatena *?verb=ListSets* a la URL base del repositorio (ejemplo: <http://tesis.repo.sld.cu/cgi/oai2?verb=ListSets>).
- *Token*: se encuesta el listado inicial de documentos concatenando *?verb=ListRecords&metadataPrefix=oai_dc* al URL del repositorio. Debe observarse al final la llave *resumptionToken* (ejemplo: http://tesis.repo.sld.cu/cgi/oai2?verb=ListRecords&metadataPrefix=oai_dc). Luego se comprueba si el token es válido. Para esto se encuesta concatenándolo en la URL de la siguiente forma: http://tesis.repo.sld.cu/cgi/oai2?verb=ListRecords&resumptionToken=archive/100/6146684/oai_dc
- *Palabras clave*: para conocer el formato en que se encuentran las palabras clave en el repositorio de manera general se encuesta el listado inicial de documentos concatenando *?verb=ListRecords&metadataPrefix=oai_dc* al URL del repositorio. Al observar el XML debe prestarse atención a los valores asociados a las llaves *dc:subject*.

La muestra de 58 fuentes en acceso abierto fue consultada entre el 15 y el 30 de marzo de 2015. En el [anexo](#) se detallan cada una de ellas desglosadas por países.

PRINCIPALES FUENTES DE PROBLEMAS

Para detectar los problemas que afectan la interoperabilidad y con ello el óptimo aprovechamiento de los fondos distribuidos de forma libre es necesario identificar los escenarios donde suelen presentarse según experiencia de los autores del trabajo. Un primer grupo surge debido a fallas en el área de las telecomunicaciones, dígame enlaces rotos, tiempos de respuesta muy grande y permisos denegados. El estándar de metadatos también es fuente de problemas. Dublin Core es el estándar empleado, este cuenta con 15 campos. El uso incorrecto de los metadatos y la omisión son los errores más comunes. Un tercer foco de problemas responde al uso del protocolo, particularmente en la respuesta a cada uno de los verbos. Los XML mal formados también son una fuente de error y se puede catalogar como un problema de corte informático.

RESULTADOS Y DISCUSIÓN

La figura 2 indica que la mayoría de los proveedores de datos usan DSpace como herramienta de software. Pero hay un por ciento significativo que son software desarrollados a la medida para cada proyecto (Otros: 34 %). En cuanto al uso de estándares de metadatos cabe resaltar que los números indican la coexistencia de varios junto con oai_dc (Fig. 3). Este aspecto no es relevante porque oai_dc está presente en todos los proveedores de datos, pero si da una muestra de la diversidad existente.

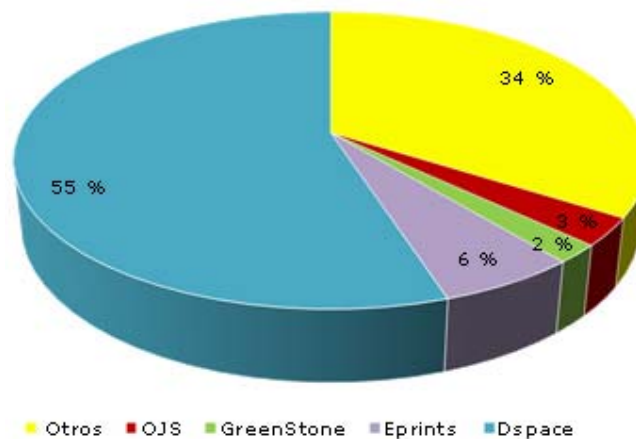


Fig. 2. Software utilizados.

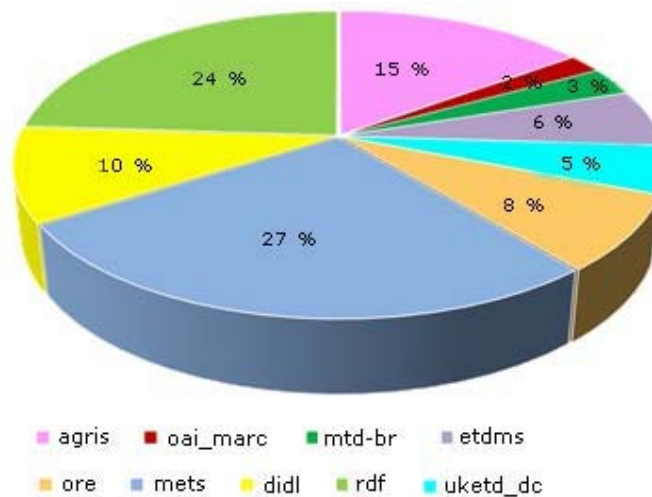


Fig. 3. Estándar de metadatos empleados, aparte de oai_dc.

El primer paso para conocer que parte de la muestra es compatible con oai-dc es identificar cuáles responden a la primera solicitud^k y cuáles no, para luego definir las causas de los posibles fallos, esto se ilustra en las [figuras 4 y 5](#). Es interesante el resultado porque evidencia que las telecomunicaciones y la informática inciden de forma importante en los fallos de los proveedores de datos, mientras que el uso del protocolo reporta una incidencia más baja.

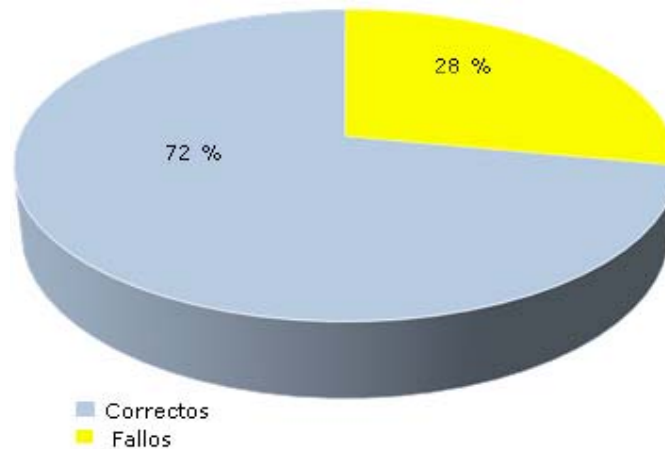


Fig. 4. Relación de respuestas correctas y fallos.

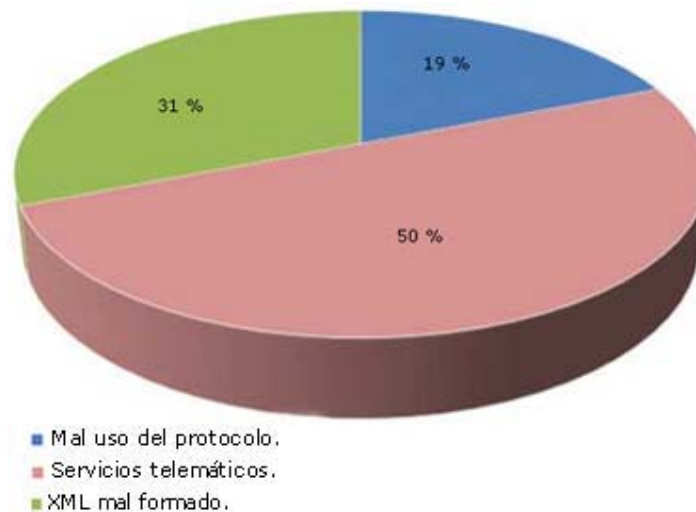


Fig. 5. Causas de los fallos.

Pero no todo termina con una respuesta correcta, porque esta puede verse afectada por un mal empleo de las normas. En la [figura 6 y 7](#) se muestran los problemas que afectaron la indexación en el experimento realizado.

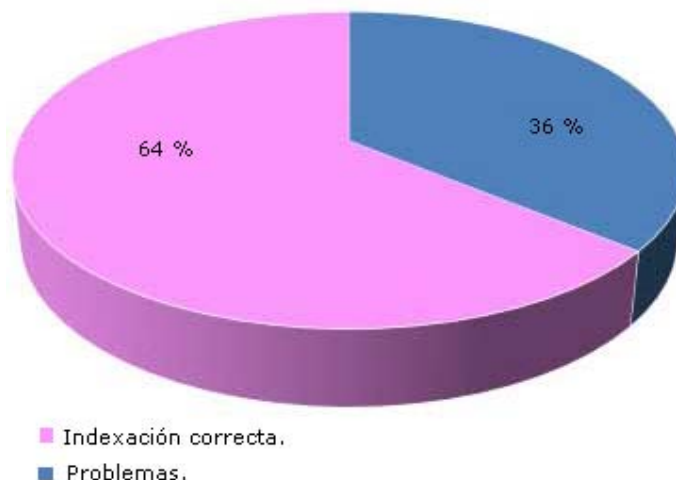


Fig. 6. Principales problemas en el proceso de indexación.

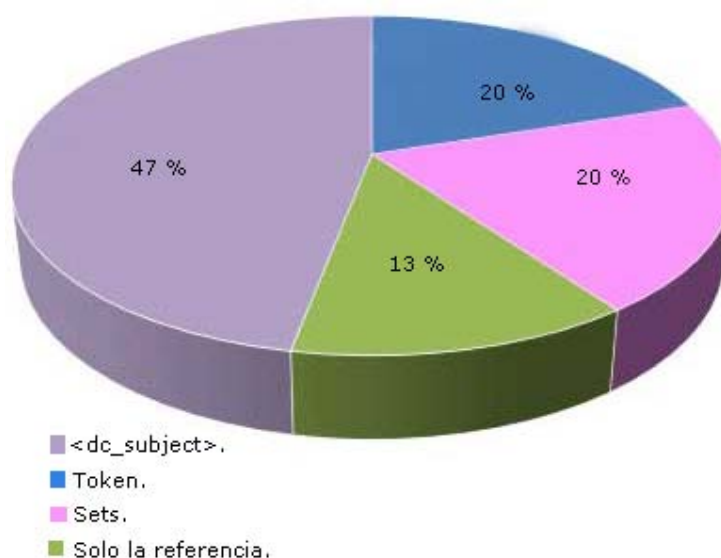


Fig. 7. Problema.

Se destacan los problemas con el uso de palabras clave, básicamente la no utilización de la etiqueta `<dc:subject>`¹ en forma de campos múltiples, usando una sola etiqueta y un carácter separador. Aunque esto no afecta el indexado, sí puede tener una incidencia negativa en la recuperación de información, sobre todo porque aumenta la complejidad de los algoritmos. Otra situación que se aprecia es el mal uso de la etiqueta, muchas veces con valores correspondientes a `<dc:type>`^m. El mal uso de `resumptionToken` y `set` también se presentan con cierta frecuencia, el

20 % de los errores encontrados corresponden a mal uso de `<resumptionToken>`. Los proveedores de datos usan en su respuesta la etiqueta `<resumptionToken>` como señal de reanudación para segmentar las respuestas cuando estas son muy grandes. Se encontraron casos en que esta etiqueta no se utiliza y esto impide la continuidad de la recolección de metadatos.

En el caso de los `set` se muestran dos dificultades: repositorios que no responden a la petición `ListSets` y otros que en la respuesta a `ListRecords` no hacen uso de la etiqueta `<setSpec>`², por lo que no se pueda asociar el registro con alguna de las colecciones del proveedor de datos y por tanto se pierde en el nivel de clasificación. También se identificaron casos en los que solo se expone el metadato referente al identificador del recurso. Esto es muy perjudicial porque no aporta suficiente información para el proceso de recuperación.

De los resultados obtenidos se puede inferir que existen demasiadas libertades en la implantación de proveedores de datos en acceso abierto, el estándar de metadatos empleado (*Dublin core*) es demasiado flexible y propicia que se produzcan errores humanos. No era objetivo de este trabajo confrontar estandarización de metadatos; sin embargo, además de los problemas identificados, se observaron irregularidades en el metadato autor (*dc:creator*) donde se pueden encontrar nombres duplicados para un mismo autor en un determinado almacén de datos, algo que se está tratando en otros trabajos buscando mecanismos para estandarizarlos. A la vista están dos posibles soluciones a largo plazo:

- El movimiento OA podría incidir en la definición del estándar (*Dublin Core*) buscando unificar el formato de los campos; por ejemplo, el campo dedicado a la fecha del documento debería tener restricciones de escritura obligando a escribir la fecha en un formato único.
- Los proveedores de servicios deben involucrarse y realizar trabajos de normalización.

CONCLUSIONES

Del estudio realizado se concluye que existe en el área de América Latina un gran número de fuentes de acceso abierto que pueden ser recuperadas y que tributan a la difusión del conocimiento. El protocolo OAI-PMH garantiza el intercambio de datos entre repositorios en acceso abiertos, sin embargo varias causas pueden afectar este proceso y de esta forma perder información de valor. Se puede identificar como tendencia en los proveedores de datos el uso de DSpace como herramienta de software, pero también se presenta un grupo significativo que hace uso de desarrollos a la medida, lo que muchas veces es causa de problemas. Los desarrollos a la medida muchas veces implican baja tolerancia a fallos; requieren de un equipo de desarrollo con alta cohesión de trabajo y experiencia para obtener productos de calidad. Los servicios telemáticos, el trabajo con los documentos XML y en la interpretación del protocolo son puntos a tener presente para proveer datos de forma eficiente y constituyen las principales fuentes de fallos. Se debe prestar especial atención a la forma de uso de las palabras clave, las colecciones (`set`) y la señal de reanudación (`resumptionToken`); varias fuentes presentan deficiencias en este sentido.

A pesar de existir un protocolo y un estándar de metadatos disponibles en pos de garantizar estandarización e interoperabilidad, persisten malos usos y malas

interpretaciones, que si bien no alcanzan valores significativos son fuentes de pérdida de valiosa información. Se recomienda que en la implementación de un proveedor de datos sea publicado el formato en que se van a almacenar las fechas, palabras clave, nombres de autor y otros metadatos para facilitar la cosecha desde proveedores de servicios. Se trabaja en un sistema de normalización de nombres de autor, el cual puede ser la base de la estandarización de los datos en proveedores de servicios.

Anexo. Fuentes de acceso abierto por países

Argentina	
Institución	Repositorios
Universidad Nacional de la Plata	http://sedici.unlp.edu.ar/ http://sedici.unlp.edu.ar/phpoai/oai2.php
Scielo Argentina	http://www.scielo.org.ar/ http://www.scielo.org.ar/oai/scielo-oai.php
Universidad Nacional de Cuyo	http://bdigital.uncu.edu.ar http://bdigital.uncu.edu.ar/OAI/index.php
Universidad Católica	http://bibliotecadigital.uca.edu.ar http://bibliotecadigital.uca.edu.ar/greenstone/cgi-bin/oaiserver.cgi
Biblioteca de Buenos Aires, Facultad de Ciencias Exactas y Naturales	http://www.digital.bl.fcen.uba.ar http://digital.bl.fcen.uba.ar/gsd1-282/cgi-bin/oaiserver.cgi
Portal de Promoción y Difusión Pública del Conocimiento Académico y Científico	http://nulan.mdp.edu.ar http://nulan.mdp.edu.ar/cgi/oai2
Brasil	
Institución	Repositorios
Scielo Brasil	http://www.scielosp.org/ http://www.scielo.br/oai/scielo-oai.php
Biblioteca Digital de tesis y disertaciones	http://bdtd.ibict.br/ http://oai.ibict.br/mypoai/oai2.php
Tribunal superior de Justicia	http://bdjur.stj.gov.br/ http://bdjur.stj.gov.br/dspace-oai/request
Universidad Federal de Rio Grande del Sur	http://www.lume.ufrgs.br/dspace-oai/request
Repositorio de Ciencias de la Comunicación	http://reposcom.portcom.intercom.org.br/ http://reposcom.portcom.intercom.org.br/dspace-oai/request

Universidad de Brasilia	http://repositorio.bce.unb.br/ http://repositorio.bce.unb.br/riunb-oai/request
Revista de Ciencias de la Información	http://www.ibict.br/cienciadainformacao/ http://www.ibict.br/cienciadainformacao/oai/index.php
Universidad Federal de Acre	http://repositorios.ufac.br:8080/repositorio/ http://repositorios.ufac.br:8080/oai/request
Universidad de Río Grande del Norte	http://www.repositorio.ufrn.br/ http://repositorio.ufrn.br:8080/oai/request
Revista de Ciencia de la Computación	http://revcom2.portcom.intercom.org.br/rbcc/ojs/ http://revcom2.portcom.intercom.org.br/rbcc/ojs/oai
Repositorio académico de Biblioteconomía y Ciencias de la Información	http://rabci.org/rabci/ http://rabci.org/rabci/oai
Revista Electrónica de Ciencias Sociales Aplicadas	http://revista.grupointegrado.br/perspectivascontemporaneas http://revista.grupointegrado.br/perspectivascontemporaneas/oai
Bolivia	
Institución	Repositorios
Escuela de posgrado, Universidad Autónoma Gabriel René Moreno	http://tesis.epg.uagrm.edu.bo/sdx/uagrm http://tesis.epg.uagrm.edu.bo/sdx/sdx/oai/uagrm/documents
Universidad Mayor San Andrés	http://www.cybertesis.umsa.bo http://www.cybertesis.umsa.bo:8080/sdx/sdx/oai/umsa/documents
Chile	
Institución	Repositorios
Scielo Chile	http://www.scielo.cl/ http://www.scielo.cl/oai/scielo-oai.php
Repositorio Académico de la Universidad de Chile	http://www.captura.uchile.cl/ http://www.captura.uchile.cl:8080/oai/request
Universidad de Talca	http://dspace.otalca.cl/ http://dspace.otalca.cl/dspace-oai/request
Universidad Autónoma de Chile	http://cybertesis.uach.cl http://cybertesis.uach.cl/sdx/sdx/oai/uach/documents
Colombia	
Institución	Repositorios
Portal de revista UN, Universidad Nacional de Colombia	http://www.revistas.unal.edu.co http://www.revistas.unal.edu.co/index.php/index/oai
Universidad del Valle	http://bibliotecadigital.univalle.edu.co/ http://bibliotecadigital.univalle.edu.co/oai/request

Scielo Colombia	http://www.scielo.org.co/ http://www.scielo.org.co/oai/scielo-oai.php
Universidad del Rosario	http://repository.urosario.edu.co http://repository.urosario.edu.co/oai/request
Colegio de estudios superiores de Administración	http://repository.cesa.edu.co/ http://repository.cesa.edu.co/oai/request
Costa Rica	
Institución	Repositorios
Scielo Costa Rica	http://www.scielo.sa.cr/ http://www.scielo.sa.cr/oai/scielo-oai.php
Universidad nacional a distancia	http://repositorio.uned.ac.cr/reuned/ http://repositorio.uned.ac.cr/oai/request
Cuba	
Institución	Repositorios
Scielo Cuba	http://www.scielo.sld.cu/ http://www.scielo.sld.cu/oai/scielo-oai.php
Repositorio de tesis de doctorados de Ciencias de la Salud	http://tesis.repo.sld.cu/ http://tesis.repo.sld.cu/cgi/oai2
Ecuador	
Institución	Repositorios
Biblioteca Flacso	http://www.flacsoandes.org/biblio/opac/index.php http://www.flacsoandes.org/biblio/oai/oai2.php
Escuela superior politécnica	http://www.dspace.espol.edu.ec/ http://www.dspace.espol.edu.ec/oai/request
Portificia Universidad católica	http://repositorio.puce.edu.ec/ http://repositorio.puce.edu.ec/oai/request
Universidad Andina	http://repositorio.uasb.edu.ec/ http://repositorio.uasb.edu.ec/oai/request
Tesis de investigaciones, Universidad de Quito	http://repositorio.usfq.edu.ec/ http://repositorio.usfq.edu.ec/oai/request
El Salvador	
Institución	Repositorios
Universidad del Salvador	http://ri.ues.edu.sv/ http://ri.ues.edu.sv/cgi/oai2

México	
Institución	Repositorios
Universidad Autónoma de México (UNAM)	http://www.rad.unam.mx http://www.rad.unam.mx/index.php/oai
Instituto tecnológico de estudios superiores de Occidente	http://www.biblio.iteso.mx/biblioteca/ http://srb2.bib.iteso.mx/oai/ite/default.aspx
Universidad de Veracruz	http://cdigital.uv.mx/ http://cdigital.uv.mx/oai/request
Redalyc	http://redalyc.uaemex.mx/ http://oai.redalyc.uaemex.mx/redalyc/oai
Documentación en ciencias de la comunicación	http://ccdoc.iteso.mx/ http://ccdoc.iteso.mx/oai.aspx
Universidad de Nuevo León	http://eprints.uanl.mx/ http://eprints.uanl.mx/cgi/oai2
Instituto de estudios superiores de occidente	http://148.201.96.14/CatIA/EDUDOCDC/ http://148.201.96.14/CatIA/EDUDOCDC/oai.aspx
Nicaragua	
Institución	Repositorios
Universidad Centroamericana	http://bjcu.uca.edu.ni:5050/dspace/ http://bjcu.uca.edu.ni:5050/dspace-oai/request
Perú	
Institución	Repositorios
Universidad Nacional Mayor de San Marcos	http://www.amauta.edu.pe http://www.amauta.edu.pe/servlet/OAI-DP
Scielo Perú	http://www.scielo.org.pe/ http://www.scielo.org.pe/oai/scielo-oai.php
Cybertesis	http://www.cybertesis.edu.pe/ http://www.cybertesis.edu.pe:80/sdx/sdx/oai/sisbib/documents
Universidad Nacional de Ingeniería	http://cybertesis.uni.edu.pe/ http://cybertesis.uni.edu.pe/sdx/sdx/oai/uni/documents
Uruguay	
Institución	Repositorios
Scielo Uruguay	http://www.scielo.edu.uy/ http://www.scielo.edu.uy/oai/scielo-oai.php
Venezuela	
Institución	Repositorios
Repositorium	http://www.repositorium.org/repositorio/ http://www.repositorium.org/repositorio/oai/

Universidad de los Andes	http://www.saber.ula.ve/ http://www.saber.ula.ve/dspace-oai/request
UCLA	http://bibadm.ucla.edu.ve/ http://bibadm.ucla.edu.ve/cqi-win/be_oai.exe
Scielo Venezuela	http://www.scielo.org.ve/ http://www.scielo.org.ve/oai/scielo-oai.php
Universidad de Zulia	http://tesis.luz.edu.ve http://tesis.luz.edu.ve/tde_oai/oai2.php

REFERENCIAS BIBLIOGRÁFICAS

1. Sancho Gómez R. Versión española de la sexta edición del Manual de Frascati. Propuesta de norma práctica para encuestas de investigación y desarrollo experimental. 2003 [citado 14 de mayo de 2013]. Disponible en: <http://redc.revistas.csic.es/index.php/redc/article/viewFile/200/255>
2. Sancho Gómez R. Indicadores bibliométricos utilizados en la evaluación de la ciencia y la tecnología. Rev Esp Docum Cient. 1990;13(3-4):842-65.
3. Díaz Pérez M. Situación de las metodologías para la medición de la ciencia, la tecnología y la innovación en América Latina. Acimed. 2009 [citado 14 de diciembre de 2014];19(4). Disponible en: <http://scielo.sld.cu/pdf/aci/v19n4/aci09409.pdf>
4. Díaz Pérez M. Producción tecnológica de América Latina con mayor visibilidad Internacional: 1996-2007. Diploma de Estudios Avanzados. Universidad de Granada, España; 2007.
5. Comisión Económica para América Latina y el Caribe. CEPAL. Anuario Estadístico de América Latina y el Caribe. 1996 [citado 2 de agosto de 2014]. Disponible en: <http://www.eclac.org/publicaciones/xml/7/35327/ANUARIO2006.pdf>
6. Albornoz Galdames M. Indicadores y política científica y tecnológica. IV Taller Iberoamericano e Interamericano de Indicadores de Ciencia y Tecnología. Red de Indicadores de Ciencia y Tecnología (RICYT); 1999.
7. Prat AM. La importancia de medir la producción científica. 2003 [citado 14 de junio de 2014]. Disponible en: <http://www.ricyt.org/interior/difusion/pubs/elc2003/8.pdf>
8. Red de Indicadores de Ciencia y Tecnología RICYT. Hacia el Manual de Buenos Aires. Indicadores de Carreras de Recursos Humanos en Ciencia y Tecnología en Iberoamérica. Buenos Aires: I Taller Iberoamericano de Indicadores de Ciencia y Tecnología; 2014 [citado 22 de abril de 2014]. Disponible en: http://ricyt.org.elsevier.com/docs/taller_rrhh/informemanualdebuenosaires.pdf
9. Albornoz Galdames M, Ratto D. Indicadores de Ciencia y tecnología en Iberoamérica. Agenda 2005. Buenos Aires: Declaración final del Sexto taller de indicadores de ciencia y tecnología iberoamericano e interamericano. RICYT; 2005.

10. Dietz JS, Chompalov I, Bozeman B, O'Neil LE, Park J. Using the curriculum vita to study the career paths of scientists and engineers: An exploratory assessment. *Scientometrics*. 2000; 49(3): 419-42.

11. Organización para la Cooperación y el Desarrollo Económicos (OCDE). Manual de Canberra. Recursos Humanos en Ciencia y Tecnología (RHCT); 1995 [citado 25 de mayo de 2013. Disponible en: <http://www.oei.es/catmexico/Manual%20Canberra.pdf>

12. D'onofrio GM, Solís Cabrera FM, Tignino VM, Cabrera E. Indicadores de trayectorias de los investigadores iberoamericanos: Avances del Manual de Buenos Aires y resultados de su validación técnica. Madrid: Observatorio Iberoamericano de la Ciencia, la Tecnología y la Sociedad del Centro de Altos Estudios Universitarios de la OEI y la Agencia Española de Cooperación Internacional para el Desarrollo (AECID); 2010.

13. Armas Peña D, Díaz Pérez M, Giraldes Reyes R. Sistema Institucional para la Gestión de la Ciencia y la Técnica en Universidades: una perspectiva cuantitativa para su análisis y evaluación. La Habana: Memorias Congreso Internacional de Información; 2008.

14. Bozeman BC, Dietz JS, Gaughan M. Scientific and technical human capital: an alternative approach to R&D evaluation. In *J Techn Manag*. 2001; 22(8): 716-40.

15. Solís Cabrera FM, Milanés Guisado Y, Navarrete Cortés J. Evaluación de la investigación científica. El caso de Andalucía. *Rev Fuentes*. 2010; 10: 83-100.

Recibido: 26 de junio de 2015.

Aprobado: 24 de noviembre de 2015.

Maikel Manuel Fernández Fernández. Universidad de las Ciencias Informáticas. La Habana, Cuba.

Correo electrónico: maikelf@gmail.com

^a Sitio Oficial OAI-PMH, <http://www.openarchives.org/OAI/openarchivesprotocol.html>

^b XML: eXtensible Markup Language. Es un lenguaje de marcado que posibilita la representación y transmisión de información.

^c HTTP: Hypertext Transfer Protocol. Es el protocolo usado en la comunicación en la World Wide Web (www).

^d ScIELO: es una base de datos bibliográfica, con artículos a texto completo. Es un modelo para la cooperación electrónica entre países. Originaria de Brasil. (<http://www.scielo.org>). Uno de los principales exponentes del movimiento de acceso abierto.

^e RedAlyC: Red de revistas científicas de América Latina, España y Portugal. Es un proyecto impulsado por la Universidad Autónoma de México con la idea de aumentar la difusión de la actividad científica en la región.

^f ROAR: por sus siglas en inglés, Registry of Open Access Repositories. Es un registro de repositorios de acceso abierto. Da información del nombre, institución encargada del mantenimiento, aspectos tecnológicos y sistematicidad de las actualizaciones. [1] ScIELO: es una base de datos bibliográfica, con artículos a texto completo. Es un modelo para la cooperación electrónica entre países. Originaria de Brasil. (<http://www.scielo.org>). Uno de los principales exponentes del movimiento de acceso abierto.

^g RedAlyC: Red de revistas científicas de América Latina, España y Portugal. Es un proyecto impulsado por la Universidad Autónoma de México con la idea de aumentar la difusión de la actividad científica en la región.

^h ROAR: por sus siglas en inglés, Registry of Open Access Repositories. Es un registro de repositorios de acceso abierto. Da información del nombre, institución encargada del mantenimiento, aspectos tecnológicos y sistematicidad de las actualizaciones.

ⁱ oai_dc: es la forma de identificar el uso del formato Dublin Core dentro de un proveedor de datos. Se retorna en la respuesta a peticiones del tipo ListMetadataFormats.

^j <resumptionToken>: es la etiqueta xml que se se usa en las respuesta de los proveedores de datos para exponer la señal de reanudación.

^k Se considera la primera solicitud, al intentar acceder a la url base del proveedor de datos.

^l <dc:subject>: es la etiqueta xml que se usa en las respuesta de los proveedores de datos para mostrar el contenido del metadato palabras clave.

^m <dc:type>: es la etiqueta xml que se se usa en las respuesta de los proveedores de datos para mostrar el contenido del metadato tipo, que se refiere al tipo de documento.

ⁿ <setSpec>: es la etiqueta xml para representar el conjunto o colección a la que está asociado un documento.