



Promover mayor visibilidad y aplicación de la investigación a través de redes globales de Repositorios de Acceso Abierto

El caso de Interoperabilidad para Repositorios de Acceso Abierto

Grupo de trabajo 2: Interoperabilidad del Repositorio

**Julio 2011
version 1.0**

<http://coar-repositories.org>

Resumen

Los Repositorios de Acceso Abierto, cuyo nombre ha ido en aumento, son un componente importante de la infraestructura de e-investigación. El valor real de los repositorios recae en el potencial de interconexión para crear una red de repositorios, una red que puede proveer un acceso unificado a los resultados de la investigación y son (re-) utilizados por las máquinas y los investigadores.

Sin embargo, para lograr este potencial, necesitamos interoperabilidad. El propósito de este documento es proporcionar un visión general de alto nivel de interoperabilidad de los Repositorios de Acceso Abierto, identificar los principales problemas y retos que deben abordarse, estimular la participación de la comunidad repositorio y poner en marcha un proceso que dirija a la creación de una hoja de ruta para la interoperabilidad del repositorio.

Acceso Abierto e Interoperabilidad

‘Acceso Abiertos’ se refiere a la práctica de conceder libre acceso a los resultados de la investigación a través de internet – de forma gratuita y libre de la mayoría de restricciones de licencia.

Por lo general las metas de Acceso Abierto son: aumentar el acceso, visibilidad e impacto de los resultados de la investigación; promover el progreso y la eficiencia de la ciencia y estimular la innovación usando la tecnología para difundir la investigación y aprobar becas, explorar, utilizar y reutilizar, y maximizar el retorno de la inversión en ciencia siendo la investigación financiada con fondos oficiales y de libre disposición al público.

Acceso Abierto se produce a través de dos métodos:

- 1) “Oro” Acceso Abierto: Los académicos publican en revistas de Acceso Abierto, es decir: revisadas por expertos, revistas especializadas que hacen que todos los artículos estén disponibles gratuitamente.
- 2) “Verde” Acceso Abierto: Los académicos publican en cualquier revista revisada por pares y luego depositan una copia de su artículo en un repositorio de Acceso Abierto.

Hoy en día, los repositorios de Acceso Abierto cada vez más están siendo utilizados para recolectar, archivar y difundir/diseminar todo tipo de resultados de investigación, tales como artículos de investigación, actas de conferencias, disertaciones, recopilados de datos, documentos de trabajo e informes.

El proceso de investigación constituye un esfuerzo internacional y distribuido, que implica una variedad de partes interesadas, tales como científicos, autores y beneficiarios de las subvenciones, centros de investigación, editoriales y agencias de financiación de la investigación. – Cada una con sus propios intereses. Las colaboraciones que se dan entre los científicos y varias instituciones, con frecuencia trabajando en países de todo el mundo. La infraestructura del repositorio imita el entorno

internacional y distribuido: La mayoría de instituciones de educación superior que están involucradas en apoyar el Acceso Abierto y mantener sus propios repositorios institucionales en los cuales ellos recolectan, archivan y difunden los resultados de la investigación de sus profesores y científicos afiliados. Algunas instituciones también mantienen repositorios amplios y especializados de disciplinas tales como arXiv.org y PubMed Central.

Cada repositorio es un valor limitado para la investigación: el poder real del Acceso Abierto recae en la posibilidad de conectarse y atarse entre repositorios, es por eso que necesitamos la interoperabilidad. Para crear una capa transparente de contenido a través de repositorios conectados alrededor del mundo, el Acceso abierto se basa en *interoperabilidad*, la capacidad para comunicar sistemas entre ellos y pasar información de ida y vuelta en un formato utilizable. La interoperabilidad nos permite aprovechar el poder computacional de hoy en día, de manera que podamos agregar, minar datos, crear nuevas herramientas y servicios, generar nuevo conocimiento del contenido del repositorio.

Estudiantes e investigadores que buscan información no necesitan saber donde fue publicado un elemento específico o donde está almacenado un artículo. En cambio, los usuarios confían en los motores de búsqueda para recuperar los artículos, y son capaces de descubrir información, de otra manera no la habrían localizado.

Por otra parte, basarse en motores de búsqueda proporciona una unión de los repositorios entre sí, el Acceso Abierto no necesita depender de un único repositorio para recoger los resultados de la investigación en el mundo. La estructura descentralizada y el valor agregado de los servicios y herramientas que se han diseñado en la parte superior de los repositorios son posibles debido a la interoperabilidad. La calidad de estos servicios depende de la información proporcionada por los repositorios y la estandarización de los datos.

La interoperabilidad es el adhesivo técnico que hace esta integración posible – y hace posibles de alcanzar las metas del Acceso Abierto.

Interoperabilidad en Pocas palabras

Dentro de los repositorios, la interoperabilidad ocurre en varias maneras. A nivel del sistema, la interoperabilidad se produce cuando los repositorios están configurados de una manera que permitan a los datos u objetos digitales pasar dentro o fuera de los repositorios a través de sistemas externos. Un ejemplo, el Protocolo para la Recolección de Metadatos (OAI-PMH por sus siglas en inglés), especifica ciertos criterios que deben cumplirse para permitir a sistemas de terceros, externos acceder y recolectar los metadatos de los repositorios.

La interoperabilidad basada en el protocolo OAI-PMH se ha convertido en parte de los estándares de implementación para repositorios. Casi todos los repositorios se adaptan a esta línea de base para la interoperabilidad y permitir que los metadatos sean recolectados. Los registros de metadatos de muchos repositorios luego son agregados y pueden ser utilizados en nuevas y diferentes formas,

muy frecuentemente en la creación de un área o portales basados en una disciplina y motores de búsqueda especializados tales como OALster y BASE (motor de búsqueda Académica de Bielefeld).

Otros protocolos también intentan aprovechar el potencial de la interoperabilidad creando sistemas que son diseñados para facilitar el intercambio de contenidos (y no solo registros de los metadatos) entre los sistemas. Tal como el protocolo iniciativa de Archivos Abiertos, reutilización e intercambio de materiales OAI-ORE (por sus siglas en inglés), que define los estándares para la descripción y el intercambio de agregaciones a los recursos de la web, algunas veces llamados como objetos digitales compuestos, permite que los elementos ya depositados en un repositorios sean copiados en otras colecciones.

Otro protocolo es, Simple Web-service Offering Repository Deposit - SWORD (por sus siglas en inglés) que permite a los autores depositar un artículo a través de una única interfaz y luego enruta el elemento a múltiples repositorios. El objetivo de SWORD es minimizar la barrera para depositar a fin de hacer el proceso tan simple y sencillo como sea posible para los investigadores. Con el fin de que cualquiera de estos protocolos se pueda trabajar, los sistemas deben estar configurados de manera consistente e interoperable.

Sin embargo, realmente este es un primer paso en interoperabilidad. Con la constante evolución tecnológica y el desarrollo de nuevas herramientas, los repositorios necesitan tener la capacidad de soportar servicios más complejos e innovadores que se basen en requerimientos adicionales para tener interoperabilidad. Gran parte del trabajo de hoy día en esta área se relaciona con otra clase de interoperabilidad: Interoperabilidad semántica. Quiere decir, asegurar que el significado preciso de la información de intercambio es comprensible y consistente en diversas formas, lo cual permite la fidelidad del significado entre las máquinas y las personas.

Este tipo de interoperabilidad permite al usuario final encontrar datos de múltiples repositorios y combinarlos de maneras significativas. Usando una terminología común de repositorio a repositorio le permite a los investigadores tener una manera más consistente para encontrar y recuperar elementos relevantes dentro de las búsquedas. Por ejemplo, usando un término permanente en varios tipos de documento permitiría a los investigadores más fácilmente encontrar los elementos que ellos buscan y podría permitir la creación de una herramienta que agrega colecciones de video de un conjunto de repositorios.

Los estándares de interoperabilidad semántica tales como el Marco de Descripción de Recursos (RDF por sus siglas en inglés) están siendo implementados para expresar las relaciones de los elementos digitales de manera que el ordenador entienda, permitiendo a las máquinas crear servicios sofisticados con representación global del conocimiento distribuido entre repositorios y otros sistemas, hacer una conexión multi-disciplinaria y combinar los resultados dispares para llegar a nuevas ideas.

Cuestiones actuales & retos con la Interoperabilidad

Mientras que la comunidad del repositorio ha desarrollado herramientas que orientan la recolección y el proceso de introducción de los contenidos, una gran cantidad de trabajo adicional aún es necesaria para abordar/dirigir otras cuestiones existentes, así como los nuevos retos originarios de la continua evolución de las e-infraestructuras globales de la investigación.

El entorno distribuido ha creado varios retos tanto a nivel técnico como en los niveles administrativo y organizacional. Algunos de los retos actuales donde las directrices de la interoperabilidad podrían conducir a una mayor consistencia en los datos de diferentes repositorios y proporcionarían apoyo a los administradores del repositorio para manejar estos desafíos de maneras consistentes son:

Retos Técnicos:

- **Nuevos tipos de contenido.** Además de los artículos de texto, nuevos tipos de contenido están siendo agregados a los repositorios de Acceso Abierto: multimedia, Conjuntos de datos y software para procesar esos datos, simulaciones, modelos computacionales, gráficos, análisis automatizados, etc. El contenido de los repositorios es cada vez más heterogéneo y dinámico, pero ninguno de los parámetros de interoperabilidad direcciona estos tipos de contenido.
- **Software y sistemas.** Mientras que la mayoría de instalaciones de los repositorios utilizan uno de muchos sistemas de software de código abierto (DSpace, Eprints, Fedora ó Invenio), los sistemas manejan procesos estandarizados en diferentes formas, conllevando a desafíos en la interoperabilidad. Además, los repositorios de Acceso Abierto son solo una parte de la e-infraestructura de investigación – los repositorios deberían ser interoperables con las revistas de Acceso Abierto, los sistemas de investigación de la información (CRIS por sus siglas en inglés), los Entornos de Investigación Virtual (VRE por sus siglas en inglés), las infraestructuras e-ciencia (ejemplo: redes computacionales), los repositorios objeto del aprendizaje, los repositorios de Recursos Educativos Abiertos (OERs), y otros tipos de colecciones digitales tales como la biblioteca de gestión de archivos digitales.
- **Nuevas capas de Servicio:** Con un incremento en los mandatos de Acceso Abierto y las políticas de la agencia de financiación de investigación a nivel institucional, y niveles nacionales, los académicos van a exigir nuevos servicios que hacen que los procesos del repositorio lleguen a ser más integrados dentro de sus flujos de trabajo. A medida que el volumen de materiales de Acceso Abierto sea mayor, los científicos también van a querer analizar y manipular el cuerpo de los datos de nuevas maneras.
- **Uso de los datos.** Para permitir una útil comparación de los sistemas, medir el impacto de las contribuciones individuales o desarrollar nuevas métricas de investigación o herramientas de monitoreo/control. Hay una necesidad de agregar e intercambiar

información de uso constante (ejemplo. Cuantas veces elementos en particular han sido descargados, visibilizados y citados)

- **Identificación consistente y terminología.** Coherencia e interoperabilidad con respecto a la identificación y denominación (autores, elementos, instituciones, agencias de financiación y subvenciones) es un requerimiento fundamental para crear espacios reales de información de los repositorios distribuidos.
- **Desafíos del Idioma.** Como el Acceso Abierto se hace más sobresaliente en todo el mundo, los retos con los lenguajes y los diferentes tipos de escritura se están volviendo más abundantes. La interoperabilidad tiene el potencial para apoyar la búsqueda y el descubrimiento a través de los servicios de traducción.

Retos Administrativos y Organizacionales

- **Contexto Global.** Actualmente, no existe ningún conjunto global de recomendaciones para presentar el contenido del repositorio o agregar contenido de otros repositorios a nivel mundial.
- **Sostenibilidad de las directrices de estandarización a largo plazo.** Algunas directrices han emergido de los proyectos de investigación, pero una vez es completado un proyecto de investigación financiado, las directrices se convierten en “huérfanas” sin una organización que asuma la responsabilidad para actualizarlas o mantenerlas. Desde que el entorno de e-investigación sea dinámico y rápidamente cambiante, es imperativo que las directrices permanezcan al día y tengan una organización responsable de su supervisión a largo plazo.
- **Apoyo a la implementación de directrices.** Como los repositorios de Acceso Abierto están cada vez más ligados a las políticas nacionales y la financiación de la investigación, será necesario para las organizaciones más pequeñas con recursos limitados implementar y mantener los repositorios. Mientras que para los miembros de la Confederación de Repositorios de Acceso Abierto (COAR por sus siglas en inglés) hay buenos ejemplos de estructuras de apoyo institucional, regional o nacional en lugares como Japón, China, América Latina y Europa. No existe una estructura como tal para apoyar la interoperabilidad de los repositorios a nivel mundial. Por lo demás, seguir las directrices y los estándares, particularmente cuando no están construidos dentro de los sistemas del repositorio, puede llegar a ser un costo restrictivo y consumir tiempo extremado puesto que las directrices cambian y evolucionan.

Los repositorios de Acceso Abierto son una pieza dinámica en la infraestructura de e-investigación que está siendo desarrollado en todo el mundo. Con el fin de descentralizar la infraestructura para un mejor apoyo a los investigadores y difundir la información de becas, los repositorios y otros sistemas deben seguir las directrices de interoperabilidad. La necesidad de interoperabilidad seguirá creciendo a medida que avanzamos hacia el futuro.

Confederación de Repositorios de Acceso Abierto (COAR por sus siglas en inglés) y la interoperabilidad

La Confederación de Repositorios de Acceso Abierto (COAR por sus siglas en inglés) ha identificado que “facilitando la discusión en interoperabilidad entre los repositorios de Acceso Abierto como parte de una infraestructura más amplia de e-investigación) como un objetivo clave para el 2010-2011. Por tal, ha sido creado y encargado un grupo de trabajo para apoyar este objetivo.

Los principales pasos previstos son.

- Recoger y analizar la información del actual entorno de interoperabilidad de las partes interesadas.
Preguntas de muestra: ¿Cuáles son los principales desafíos no contemplados en las actuales directrices de la interoperabilidad? ¿Qué problemas existen para la comunidad de investigación que podrían ser abordados por la interoperabilidad en el futuro? ¿Qué problemas relacionados con la interoperabilidad actúan como barreras hacia el progreso? ¿Cuál debería ser el ámbito para las directrices emitidas, apoyadas y/o mantenidas por la Confederación de Repositorios de Acceso Abierto (COAR por sus siglas en inglés)?
- Dirigir una evaluación ambiental. Repasar las directrices actuales, anticuadas y huérfanas, las directrices en desarrollo, y los cambios tecnológicos que podrían tener impacto en la interoperabilidad en un futuro cercano.
- Establecer una hoja de ruta de la Confederación de Repositorios de Acceso Abierto (COAR por sus siglas en inglés) para la interoperabilidad del repositorio basada en la información de las partes interesadas, intentando maximizar el (re)uso de las directrices existentes y las mejores prácticas.

Todos los interesados están invitados a participar en los meses venideros a dar información para identificar las necesidades y los desafíos existentes, en trabajo con la Confederación de Repositorios de Acceso Abierto (COAR por sus siglas en inglés) para ayudar a desarrollar la hoja de ruta, discutir la necesidad de las directrices para la COAR y participar en el proceso. Esperamos contar con su participación.

Reconocimientos y contribuyentes

Editores

- Eloy Rodrigues (Universidade do Minho, Portugal)
- Abby Clobridge (Clobridge Consulting, United States)

Expertos y Revisores

- Theo Andrew (EDINA, University of Edinburgh, United Kingdom)
- Andi Aschenbrenner (Goettingen State and University Library, Germany)
- Magchiel Bijsterbosch (SURF, The Netherlands)
- Stefan Buddenbohm (Goettingen State and University Library, Germany)
- Peter Burnhill (EDINA, University of Edinburgh, United Kingdom)
- Leonardo Candela (CNR, Italy)
- Pablo de Castro (Universidad Carlos II de Madrid, Spain)
- Donatella Castelli (CNR, Italy)
- Neil Jacobs (JISC, United Kingdom)
- Alicia López Medina (UNED, Spain)
- Norbert Lossau (Goettingen State and University Library, Germany)
- Anja Oberländer (University of Konstanz, Germany)
- Robin Rice (EDINA, University of Edinburgh, United Kingdom)
- Birgit Schmidt (Goettingen State and University Library, Germany)
- Friedrich Summann (University Bielefeld, Germany)
- Alma Swan (Key Perspectives Ltd., United Kingdom)

Traducción al español

- Malgorzata Lisowska (Universidad del Rosario, Colombia)

Anexo I: Recursos para una mayor información sobre Acceso Abierto

[A Very Brief Introduction to Open Access](#) by Peter Suber or Peter Suber's [Open Access Overview](#)

[Budapest Open Access Initiative](#) (2001) – One of the original statements in support of Open Access, now an open document that can be signed by the public.

[Berlin Declaration of Open Access to Scientific Knowledge](#) (2003) – A major international statement in support of Open Access. Organizations that sign the document commit to implementing this definition of Open Access.

[Directory of Open Access Repositories \(OpenDOAR\)](#) – A directory of academic Open Access repositories.

[Open Access Scholarly Information Sourcebook \(OASIS\)](#) – Practical steps for implementing Open Access.

[SHERPA JULIET](#) – Summaries of funding agencies' grant conditions on self-archiving of research publications and data.

[Registry of Open Access Repositories \(ROAR\)](#) – Information about the growth and status of repositories throughout the world.

[Registry of Open Access Repository Material Archiving Policies \(ROARMAP\)](#) – Links to institutional and funder mandates and repositories throughout the world.

Anexo II: Pautas clave de Interoperabilidad, Protocolo, Proyectos y Recursos

[DINI Certificate for Document and Publication Services](#) (2010) – un certificado que describe los aspectos técnicos, organizacionales y legales (incluyendo la interoperabilidad) que deben ser considerados en la creación de un servicio para Repositorios Académicos.

[DL.org Community](#) – Interoperabilidad de la biblioteca digital, mejores prácticas y modelos fundados.

[DRIVER Project](#) – Un proyecto de dos fases que establece tanto la infraestructura técnica como organizativa de una red de repositorios en Europa, y produjo varios estudios ([several studies](#)) y Directrices DRIVER ([DRIVER Guidelines](#)).

[euroCRIS](#) – La Organización Europea para la Información Internacional de Investigaciones ha estado trabajando en la interoperabilidad entre CRISs ([Current Research Information Systems](#)), diseñado para almacenar datos sobre las actividades de investigación actuales, becas, investigadores, etc. y Repositorios Institucionales.

[Knowledge Exchange Usage Statistics Guidelines](#) – Directrices para la agregación y el intercambio de los datos.

[Open Archives Initiative – Object Reuse and Exchange \(OAI-ORE\)](#) – define estándares para la agregación de objetos digitales compuestos.

[Open Archives Initiative – Protocol for Metadata Harvesting \(OAI-PMH\)](#) – protocolo por el cual los repositorios permiten que sus metadatos sean cosechados y recogidos por otros sistemas.

[Publishing and the Ecology of European Research \(PEER\)](#) – proyecto para “investigar los efectos de gran escala, el depósito sistemático de los autores finales de los manuscritos revisados por pares.

[PersID](#) – Proyecto para construir un identificador de una infraestructura persistente para las publicaciones digitales y los recursos electrónicos.

[Resource Description Framework](#) – Un modelo estándar para el intercambio de datos en la web.

[Scholarly Output Notification and Exchange \(SONEX\)](#) – Una iniciativa apoyada en JISC para identificar y analizar fasos útiles para el depósito de documentos de investigación en los repositorios.

About COAR

La Confederación de Repositorios de Acceso Abierto (COAR por sus siglas en inglés) , es una asociación reciente de las iniciativas del repositorio que comenzó en Octubre de 2009, uniendo más de 80 miembros y socios de 24 países toda Europa, Latino América, Asia y Norte América. Su misión es aumentar la visibilidad y aplicación de los resultados de la investigación a través de las redes globales de Repositorios Digitales de Acceso Abierto.

Más información acerca de la Confederación de Repositorios de Acceso Abierto (COAR por sus siglas en inglés) y sus miembros, está disponible en el sitio web de la COAR:

<http://www.coar-repositories.org> .

Información de contacto:

COAR Office
Dr. Birgit Schmidt
Katharina Müller
c/o Goettingen State and Univesity Library
Platz der Gottinger Sieben I
37073 Goettingen
Germany
Phone: +49 551 39-22215
Fax: +49 551 39-5222
coar-office@sub.uni-goettingen.de