

## **Estudio de la integración de repositorios en el sistema científico-investigador: alternativas y estado actual**

**Autores:** Lorenzo Gil, Emilio; Braña Ferreiro, Eva; Nieto Caramés, Sergio

**Palabras Claves:** Repositorios, Interoperabilidad

### **Resumen:**

En la actualidad, la escasa aceptación de los repositorios académicos de investigación en la comunidad universitaria supone uno de los principales desafíos de trabajo interno institucional. Incluso cuando hay mandatos de depósito, los repositorios se perciben como nichos aislados en el sistema científico-investigador, y posiblemente se siguen considerando como una burocracia adicional en los procesos de difusión de la investigación y acreditación de publicaciones, método y procesos de trabajo de los investigadores, etc.

Para cambiar la percepción antes apuntada, se requiere rediseñar el concepto de repositorio para integrarlo mejor con las herramientas y procesos de los investigadores, proporcionándoles funcionalidades ajustadas a sus necesidades y métodos de trabajo y comunicación académica. De este modo, los repositorios podrán ser aceptados y ser considerados herramientas que aportan valor en el corto y medio plazo a la comunidad investigadora, mas allá, claro está, de los inconvenientes derivados del aprendizaje y uso de estos sistemas. El reto, como puede advertirse, no es otro que convertirlos en herramientas centrales de la comunicación y producción científico-académicas.

Consideramos que el rediseño de los repositorios que debe ir unido a su integración real y efectiva con el resto del ecosistema científico-investigador. Éste, constituido por un panorama creciente de sistemas dedicados a cubrir una amplia variedad de elementos y funcionalidades, por otro lado tanto internos como externos, están evolucionando permanentemente de acuerdo a los requisitos de la comunidad investigadora evidenciando la urgente necesidad de adaptación, igualmente continua, de sus repositorios. Percibimos que de forma paulatina, cada vez más repositorios de instituciones universitarias se conciben o evolucionan hacia modelos integrados, haciendo uso de la mayor capacidad de integración que se construye tanto en los software base de los repositorios, como en aplicaciones y servicios potencialmente interoperables.

En la primera parte del estudio realizado se analizan la variedad de enfoques, tecnologías y procedimientos que se pueden emplear en los repositorios para lograr esa integración y complemento funcional de los repositorios. Pertenecientes al ecosistema conectado podemos relacionar una larga lista de sistemas como *Sistemas de recolección* pertenecientes a infraestructuras nacionales e internacionales, *Sistemas de Gestión de Investigación*, *Sistemas de Gestión de Bibliotecas*, *Sistemas de identificadores de autor*, *Repositorios cooepitadores*, *Servicios de información de derechos de edición*, *Gestores de referencias bibliográficas* y *Sistemas de métricas de impacto y altmétricas*, entre otros.

Consideramos que hay un gran potencial de desarrollo de los repositorios, y de su aceptación como componentes esenciales de la infraestructura de científica-investigadora, si éstos incorporan a su diseño estos modelos integrados. De hecho, la integración con

algunos de los anteriores servicios, como la disponibilidad de medidas de impacto y altmetrías, se ha señalado como un incentivo para la adopción de los repositorios por los investigadores, en concreto para incrementar las tasas de autodepósito.

En una segunda parte del estudio, efectuamos un análisis cuantitativo de la presencia de estas integraciones en los repositorios universitarios de investigación españoles. El estudio *front-office* se ha centrado, aunque no limitado, en la medición del subconjunto de integraciones existentes, que bien podríamos señalar como visibles o *exteriores*, dadas las dificultades asociadas a la observación de las integraciones que se ofrecen en procesos internos de los repositorios, como pueden ser los autodepósitos o integraciones CRIS.

El resultado final, además de caracterizar el estado del arte de la integración de los repositorios digitales, dibuja potencialidades de evolución para que los responsables de estos sistemas puedan planificar el mapa de alternativas posibles para convertirlos en puntos centralizados para el depósito, difusión y preservación del material académico y científico de las instituciones universitarias.

## INTRODUCCION

Se lleva tiempo señalando la importancia creciente de los repositorios y sus servicios asociados como componentes importantes de la infraestructura de investigación. Como apuntaba el informe *Estado de la interoperabilidad de repositorios* (Confederation of Open Access Repositories, COAR, 2012), la utilidad real de los repositorios reside en su potencial para formar parte de redes interconectadas de repositorios, aunque como este mismo informe continuaba, la infraestructura de repositorios evolucionaba rápidamente hacia un panorama de interoperabilidad descrito como "*caótico, confuso y complejo*".

Este mismo año 2015, el informe *Direcciones futuras en la interoperabilidad de los repositorios*, (Confederation of Open Access Repositories, COAR, 2015), recomienda que para evitar que los repositorios se conviertan en silos, se debe asegurar la interoperabilidad de sus contenidos y sistemas.

Como consideración adicional, existe la necesidad de comprender cómo la interoperabilidad debe incluirse como criterio básico del diseño de los repositorios, pues éstos ya no existe únicamente conectados con la estructura de agregadores de ciencia nacional e internacional, sino que cada vez coexisten e interactúan con sistemas de la propia institución, como sistemas de e-learning, sistemas de gestión bibliotecaria, metabuscadores, etc.,, y con sistemas externos, como plataformas de editores, sistemas de preservación..

La interoperabilidad en este amplio contexto puede interpretarse como la medida de la capacidad de sistemas, aplicaciones y datos diferentes unos de otros, de compartir información y funcionalidades de un modo correcto y proporcionando más valor que un único sistema, en el que todos ellos se encontrarán integrados (Gordon, 2003)

La interoperabilidad de los repositorios permite agregar contenidos, explotar contenidos de otros repositorios, ofrecer nuevas herramientas y servicios a sus usuarios, generar y aportar nuevos conocimientos a los investigadores sobre sus publicaciones y en suma,

nuevas posibilidades para integrarse fuertemente con los procesos de investigación y publicación. El repositorio debe demostrar que proporciona beneficios a corto y medio plazo adaptándose a las prácticas de trabajo de los investigadores de cada institución (Bell & Sarr, 2010)

La interoperabilidad se considera de forma habitual como uno de los factores de evaluación de los repositorios (Serrano *et al*, 2014), y de forma específica, el grado o nivel de interoperabilidad sirve para medir la capacidad de contribución del repositorio al conocimiento académico nacional e internacional (Westell, 2006) o su adaptación y aportación a cubrir las necesidades de una institución y de sus investigadores (Thibodeau, 2007). El estudio de las posibilidades técnicas de interoperación de los repositorios cobra una especial importancia para los responsables de estas infraestructuras.

En la primera parte de este trabajo, se proporciona un revisión detallada de las posibilidades actuales de integración de los repositorios, analizando la variedad de enfoques, tecnologías y procedimientos que se pueden emplear para lograr esa integración y complemento funcional. Al mismo tiempo, se intenta aportar claves para entender la complejidad e importancia de cada área de interoperabilidad

Como integrantes del ecosistema investigador hemos incluido, aunque la lista de antemano asumimos que es limitada, distintos ejes de estudio: Sistemas de agregación, Sistemas de gestión de investigación, Sistemas de gestión bibliotecaria, Sistemas de identificadores de autor, Repositorios cooperativos, Servicios de información de derechos de edición, Gestores de referencias bibliográficas, Sistemas de estadísticas, y Sistemas de métricas de impacto y altmétricas.

## **Sistemas de Agregación**

Posiblemente el modo de interoperabilidad mas conocido de los repositorios es el basado en el protocolo OAI-PMH, cuyo funcionamiento y uso en el contexto de los repositorios se ha descrito extensamente (Barrueco *et al.*, 2003). Al permitir la recolección y agregación de metadatos de repositorios, forma la espina dorsal de las infraestructuras de difusión de los sistemas de ciencia y tecnología (Recolecta, LA Referencia, REMERI, OpenAire,..) o de agregadores temáticos (Agris, OceanDocs..). Como mayor ventaja de este enfoque de integración señalaríamos la amplia difusión y estabilidad del protocolo OAI-PMH, presente en multitud de sistemas, entre ellos la totalidad de los sistemas de repositorios, agregadores, sistema de publicaciones electrónicas, sistemas de gestión bibliotecaria, metabuscadores, etc.. Como inconvenientes, señalar la necesaria normalización de metadatos (p. ej. Driver) para una integración efectiva y la poca riqueza descriptiva de los esquemas de metadatos interoperados, generalmente Dublin-Core simple.

Se han propuesto e implementado extensiones al protocolo OAI-PMH, como OAI-ORE (Open Archives Initiative, s.f.) y otros sistemas de agregación de recursos (Van de Sompel *et al.*, 2004) para permitir la recolección de objetos completos (texto completo) aunque su uso no parece estar generalizado. Igualmente importante, aunque todavía por evaluar su aplicación real en futuros escenarios de integración, es la RDF-ización de los contenidos de los repositorios disponible en versiones recientes.

## Sistemas de Gestión de la Investigación

Los sistemas Current Research Information System, CRIS, son los sistemas de información que gestionan y almacenan la información sobre la investigación realizada por una Institución. Dado que las publicaciones son una parte importante de la investigación, hay un creciente interés en la necesidad de conectar los sistemas CRIS con los repositorios, considerados estos como los mecanismos de preservación y difusión de esos resultados finales del proceso investigador.

Dada la amplia variedad de sistemas CRIS existentes, en cobertura o alcance funcional y en tecnologías usadas, no parece existir un modo predominante o estándar de interconexión entre repositorios y CRIS. De hecho, hemos inventariado esquemas de integración dirigidos desde los CRIS institucionales (el depósito se realiza en origen CRIS y de ahí se deposita en el repositorio) y desde los repositorios (los objetos se validan y enriquecen en el repositorio y de ahí se integra en el CRIS), usando una amplia variedad de enfoques tecnológicos.

No obstante, podemos identificar dos mecanismos principales para efectuar la integración CRIS-Repositorios: protocolos SWORD y protocolos REST.

El uso del protocolo SWORD (Allinson *et al.*, 2008) puede considerarse un estándar de interoperabilidad orientado al depósito en repositorios desde múltiples fuentes y formatos. De historia breve (se define en el año 2007) se está usando en una variedad de sistemas CRIS, conectando los pasos finales del proceso de investigación con los repositorios institucionales. No obstante aclararíamos que su uso no debe considerarse limitado al depósito desde sistemas CRIS sino que puede usarse (Lewis *et al.*, 2012) y de hecho, se usa, por una variedad de sistemas con interfaces de depósito ajenas o que complementan o extienden la funcionalidad de depósito de los repositorios como mencionaremos más adelante.

El segundo escenario de esta integración sistémica hace uso de arquitecturas o protocolos REST, Representational State Transfer, para construir interfaces no estrictamente limitados al depósito. Mediante un conjunto razonablemente completo de puntos de acceso REST al repositorio, abarcando funciones de creación-recuperación, actualización y borrado, un sistema externo como un CRIS puede construir su propia interface de integración sistémica, abarcando aspectos de búsqueda y difusión de objetos, gestión de usuarios, realización de depósitos, gestión de la preservación, etc. siendo factible el acceso a la mayoría de las funciones OAIS haciendo uso de este protocolo.

Por las características anteriores, se visiona este modo de integración como el modo más flexible en la construcción de servicios de interoperabilidad o integración, habiéndose documentado hasta la fecha integraciones (además de las realizadas con sistemas CRIS) con sistemas de aprendizaje LMS, como Moodle y Blackboard, y con sistemas de gestión de contenidos CMS, como Joomla y Drupal. Siendo el interface REST una manera efectiva de acceder, re-usar y gestionar contenidos almacenados por repositorios desde otros sistemas, posiblemente asistamos a una eclosión de su uso en el medio plazo.

Señalar que los dos modelos anteriores, posiblemente derivado de su concreción tecnológica un poco tardía, coexisten con otros enfoques o modelos de interoperabilidad, habiendo identificado a) archivo basado en la creación de paquetes SIP, Submission

Information Package; b) integraciones directas con la capa de datos del repositorio, es decir accesos directos a la BBDD; c) recolección OAI-PMH por el sistema CRIS; d) reconstrucción de los eventos de actualización de los objetos en el sistema destino; entre otros.

## **Sistemas de Gestión de Bibliotecas**

Como fuente de los datos bibliográficos de las publicaciones los Sistemas de gestión de Bibliotecas, LMS, Library Management Systems, son un candidato obvio para la realización de integraciones con los repositorios, aunque la fuerte presencia de sistemas propietarios en este ámbito parece haber retrasado la adopción de integraciones fuertes entre ambos mundos. Se presentan un par de opciones de construcción de la integración, pudiendo el repositorio consultar y extraer contenido del sistema bibliotecario o realizando la exportación de contenidos del repositorio hacia la catalogación en el sistema bibliotecario.

Entre las posibilidades de integración, la disponibilidad general del interfaz Z39.50 en los sistemas bibliotecarios le hace un primer candidato para la consulta del catálogo por el repositorio, siendo una segunda opción la construcción de pasarelas o crosswalks de metadatos específicas para la incorporación de registros bibliográficos al repositorio.

Igualmente factible es la utilización del protocolo OAI-PMH para la realización de la integración inversa, la actualización de los registros en el sistema de gestión bibliotecario, enfoque de interoperabilidad que comparten la práctica totalidad de los portales de descubrimiento y metabuscadores, (Vufind, Primo, Summon,...) por mencionar sistemas cercanos o complementarios a los LMS.

## **Sistemas de Identificadores de Autor**

Los sistemas de identificación unívoca de autores pretenden dar una solución al problema de la identificación y desambiguación de los autores. Existen un número considerable de sistemas y proyectos que intentan abordar esta cuestión, desde una perspectiva nacional, disciplinar, comercial, etc.. (VIAF, ORCID, ResearcherID, Google Scholar, ScopusID..)

El primer aspecto del problema es la selección de identificadores, señalado por Borrrego (Borrrego, 2013) y derivado de la abundancia de propuestas de finalidad similar, que han ido dejando su impronta en la comunidad investigadora, con diferentes tasas de adopción o penetración. De esta forma, hay que considerar la conveniencia de incorporar múltiples identificadores, precisamente para lograr el enlace, es decir interoperabilidad, con sistemas diversos que recogen la producción de los investigadores.

Añadiríamos que en el contexto de los repositorios universitarios (recordemos que los repositorios de investigación de las instituciones universitarias españolas, son, además y mayoritariamente, repositorios institucionales), aparece la necesidad de incorporar identificadores de autores incluidos como autoridades de sistemas nacionales como VIAF, con el fin de construir modelos RDF necesarios para la compatibilidad con el *Europeana Data Model*.

En el caso de realizarse esta conexión con los distintos identificadores de autor, (hay otros enfoques, como realizar este enlace en el sistema CRIS) el punto de integración se realiza principalmente a partir de la funcionalidad de autoridades del repositorio. La desambiguación asociada a los metadatos de autor es la puerta de acceso a la interoperabilidad con los servicios ofrecidos por las plataformas que operan estos sistemas, desde los meros servicios de interconexión visual en forma de enlaces desde el repositorio a los perfiles de autor existentes en Google Scholar, por ejemplo, a los servicios de integración sistémica con sincronización de publicaciones, biografías, filiaciones y resto de datos relacionados con el perfil de autor, que p. ej. ofrece ORCID mediante su REST-API (ORCID, s.f.) pasando por integraciones manuales usando las capacidades de exportación-importación de referencias bibliográficas, p. ej. a través de BibTex, que se dispone en algunos de estos sistemas.

### **Repositorios Coopetidores (cooperando-compitiendo)**

Dada la importancia que en determinadas comunidades científicas tienen los repositorios disciplinares o temáticos, (arXiv, PLOS, PubMed, Zenodo...), su uso es prioritario en estos colectivos, depositándose incluso en estos repositorios antes de su publicación en cualquier otro medio. Igualmente importante es el interés de las Instituciones investigadoras de simplificar el depósito a los investigadores, minimizando el problema del doble-depósito, maximizar la exposición de sus trabajos y capturar todo el contenido relevante posible.

Así, algunos de estos repositorios, como arXiv (Cornell University Library, s.f.), o Zenodo (Zenodo, s.f.) ofrecen acceso a través de interfaces programáticas, APIs, a los contenidos y metadatos almacenados, pero también al interfaz de depósito de publicaciones, posibilitando de esta manera la integración efectiva con estos sistemas y nuestro repositorio.

La interoperabilidad se puede construir desde ambos extremos, bien siendo el repositorio institucional el que extraiga contenido del repositorio temático, bien exportando contenidos hacia este último, en una suerte de modelo de cooperación-competición que posiblemente nos resulte familiar en el futuro.

En la misma línea determinados repositorios, como por ejemplo BioMed Central, ofrecen servicios denominados de depósito directo o automatizado (BioMed Central Ltd., s.f.), consistente en el depósito mediante SWORD de las publicaciones de una institución, después de que el autor haya finalizado el depósito en el repositorio BioMed.

De forma similar, aunque no consideremos a los sistemas de información de los grandes grupos editoriales como repositorios, encontramos servicios que se ofrecen a través de interfaces programáticas similares a los descritos anteriormente. La API-REST de Elsevier (Elsevier, s.f. -a) permite acceder a los datos de Scopus para extraer los metadatos de publicaciones de los sistemas de Elsevier e incorporarlos a los sistemas destino, los repositorios.

### **Servicios de información de derechos de edición**

Como aspecto principal del proceso de depósito, la mayoría de repositorios recalcan la importancia de la comprobación, para trabajos ya publicados, de las condiciones en que se has cedido determinados derechos a los editores.

Dicha comprobación se puede realizar mediante la herramientas SHERPA/RoMEO (SHERPA/RoMEO, s.f. -a) y DULCINEA (DULCINEA, s.f.) proyectos ambos que, en un ámbito internacional y nacional respectivamente, recopilan las políticas editoriales.

Desde el punto de vista que nos ocupa, la integración con SHERPA/RoMEO se realiza mediante la conexión a la correspondiente API (SHERPA/RoMEO, s.f.-b), mientras que al no disponer el proyecto DULCINEA de un punto de entrada programático, la posibilidad más directa pasa por el uso de técnicas de *data-scraping*, extrayendo los datos pertinentes de las respuesta de este aplicativo a las consultas realizadas en su interfaz.

### **Gestores de Referencias bibliográficas**

El modo de integración principal es la exportación e inclusión de la referencia bibliográfica de un objeto del repositorio a alguno de los gestores de referencias bibliográficas de uso predominante en las bibliotecas, RefWorks, EndNote, ProCite y BibTex entre otros, formatos de intercambio como RIS o los nuevos sistemas de recolección de citas como Zotero o de recogida y compartición como Mendeley.

En su forma más general, esta integración se realiza mediante programas que analizan el tipo de objeto a exportar, a continuación mapean los metadatos existentes en el repositorio a los campos de citación correspondientes y por último realizan la inclusión de la exportación en el gestor de referencias. Sus requisitos de funcionamiento son variados, una adecuada configuración de la conexión con el gestor, como RefWorks, o el uso embebido de OpenURL CoinS o RDF embebido, en el caso de Zotero

La integración con estos sistemas o plataforma se logra con una diversidad de enfoques, en ocasiones incorporando al repositorio código simple RIS, en otras, incorporando plugins o, código adicional de una determinada distribución del software del repositorio caso de RefWorks, en otras como extensiones de los navegadores, Zotero, o requiriendo del usuario la conexión a aplicaciones web específicas, como Mendeley.

### **Sistemas de Estadísticas de uso del repositorio**

El uso efectivo de un repositorio debe ir acompañado de la medición de su uso, actuando las estadísticas como un fuerte incentivo para la contribución de los investigadores al repositorio (Confederation of Open Access Repositories, COAR, 2013) Aunque es simple saber cuantos objetos se han enviado o almacenado, es un poco más difícil saber cómo y cuánto se han usado esos objetos, en términos de acceso a los metadatos descriptivos, descarga y visualización de artículos, origen del tráfico, etc.

Sin entrar en la precisión de los diferentes sub-sistemas de estadísticas que a lo largo del tiempo se han ido incorporando o adjuntando a los repositorios como añadidos, add-ons, ya sean código abierto, sistemas propietarios, desarrollos a medida, señalaríamos que dichos elementos tienen diferentes interpretaciones sobre los términos “descarga”,

“acceso”, “visualización” o “ítem” por lo que la comparación de valores se ha visto siempre cuestionada.

El registro del tráfico del repositorio mediante sistemas externos ha sido posible mediante dos mecanismos principales, la recogida de datos por Google Analytics o la recogida por sistemas frontales al repositorio, interceptando los datos de tráfico en el servidor web y procesando dichos datos por software específico de análisis de tráfico web, como Awstats (AWStats, s. f.) o Sawmill (Flowerfire, Inc., s. f.) entre otros.

Por otra parte la visualización de esos mismos datos asociado a los objetos del repositorio, importante como valor añadido a las publicaciones y autores, ha sido posible únicamente en recientes versiones del software (DURASPACE, s. f.) para las estadísticas de Google Analytics.

### **Sistemas de métricas de impacto y altmétricas**

Las métricas de impacto, como el factor de impacto y el h-index, se han usado tradicionalmente como dos de los principales indicadores usados en la cuantificación de la influencia de la actividad investigadora y académica. La premisa básica de las métricas basadas en citas es que el número de citas está directamente correlado con el impacto académico de publicaciones, artículos y autores.

Diversos actores, ligados principalmente a los principales grupos editoriales, ofrecen el acceso a determinadas métricas de impacto por medio de mecanismos de integración basados en APIs. Las condiciones de acceso y uso, así como los resultados, son diversos, por ejemplo la API de Cited By Scopus (Elsevier, s.f. -b) ofrece las citas por medio de una búsqueda usando los identificadores del artículo. Los resultados se muestran como una imagen con la marca Scopus que puede ser embebida en el repositorio.

Las altmétricas o altmetrías miden el impacto de la investigación mediante el uso de métricas alternativas a las métricas de citas, cuantificando su presencia en las redes sociales, en forma de menciones en la web social, descargas, enlaces, cobertura mediática, inclusión en gestores de referencias, etc. Aunque su uso para la evaluación científica está aún en proceso de estudio (Borrego, 2014), es innegable la eclosión de sistemas y aplicaciones que pueden usarse como complemento de la información sobre un determinado artículo o autor.

El número de proveedores de este tipo de servicios es cambiante, desapareciendo o integrándose servicios, apareciendo nuevos actores producto de adquisiciones corporativas, etc. Con una estabilidad relativa, dentro de la relativa novedad de estos indicadores, mencionaríamos a Almetric y a PlumX.

La incorporación de estos nuevos indicadores a nuestro repositorio se hace generalmente mediante mecanismos de API, siendo de uso libre, de momento, la correspondiente a altmetric (Altmetric, s.f.) y de uso “comercial” la API proporcionada por plumX (Plum Analytics, Inc., s.f.). Reseñar que la integración con PlumX también se ofrece mediante widgets que integran visualmente los resultados en nuestro repositorio. Ambos mecanismos permiten una integración rápida de las altmétricas de artículo en un



repositorio, aunque nos queda el interrogante de la volatilidad de estas soluciones técnicas ofrecidas en un escenario inestable como el actual.

## **ESTUDIO CUANTITATIVO: METODOLOGÍA**

Adicionalmente consideramos de interés caracterizar la presencia de estas integraciones en los repositorios universitarios de investigación españoles, con el objetivo de tener una visión general del uso efectivo de estas integraciones y del momento de adaptación de los repositorios al panorama de interoperabilidad descrito. Por ello, en la esta segunda parte del estudio, efectuamos un estudio cuantitativo de las integraciones existentes, que podríamos señalar como visibles o exteriores, dadas las dificultades al observar las integraciones que se ofrecen en proceso internos de los repositorios, como pueden ser los autodepósitos o integraciones CRIS.

En el presente estudio se han medido los repositorios universitarios de investigación españoles que aparecen recogidos en el Directorio de repositorios institucionales de Rebiun (Red de Bibliotecas Universitarias, REBIUN, s.f.) y además figuren en el Ranking Web de Repositorios (Cybermetrics Lab; Consejo Superior de Investigaciones Científicas, CSIC, s.f.), resultando un universo de 57 repositorios

Las medidas se realizaron en la semana de 16 al 22 de febrero de 2015, siendo el sistema de medida empleado la navegación como usuario anónimo por la interfaz de usuario de los repositorios estudiados. La limitaciones de no poder acceder a la interface como usuario autenticado o como usuario de autoarchivo, necesarias por otro lado para complementar la medición, han sido complementadas por la lectura de las guías y páginas de ayuda de cada repositorio, con el fin de completar la lista de integraciones disponibles.

Así, se ha podido medir la presencia de una serie de integraciones en su mayor parte visibles o externas, tales como identificadores de autor, herramientas de exportación-importación de referencias bibliográficas, sistemas de métricas de impacto y altmétricas etc.

## **ESTUDIO CUANTITATIVO: RESULTADOS**

### **Sistemas de Agregación**

El nivel básico de interoperabilidad de los repositorios con los sistemas de agregación/recolección, es conseguido por la casi totalidad de los repositorios, 55 repositorios de los 57 observados, aunque un 54% es reportado por Recolecta con problemas de compatibilidad Driver y un 39 % sin compatibilidad OpenAire (FECYT, s. f.)

### **Sistemas de Gestión de la Investigación**

Al menos tres (5%) de los repositorios analizados tienen capacidades de realizar el

depósito desde el sistema de Gestión de Investigación al repositorio mediante el uso de interfaces SWORD. Dos adicionales realizan la integración CRIS-repositorio mediante la creación de paquetes SIP y posterior ingesta de esos paquetes.

### Identificadores de Autor

Un pequeño número de repositorios, diez, se integra débilmente, integración mediante hipervínculos desde autores, con sistemas de identificadores de autor, aunque es reseñable que dos de ellos enlacen hasta 6 posibles identificadores diferentes. Además Google Scholar ID (habría que hablar de Google Scholar Citation Profile realmente) es de uso preferente sobre otros identificadores.

Repositorios con Integración	47	5	2	1	2
Nº de identificadores soportados	0	1	2	3	6

Tabla 1: Nº de Identificadores soportados

Viaf	ScopusID	Dialnet	ORCID	ResearcherID	Google Scholar
2	3	3	4	4	8

Tabla 2: Nº de repositorios integrando un determinado Identificador

### Repositorios Coopetidores (cooperando-compitiendo)

No se han identificado, aunque aclaremos que la metodología de medida no lo facilita, integraciones de los repositorios analizados con otros repositorios, los denominados coopetidores.

### Servicios de información de derechos de edición

Una información habitual en los repositorios es la relativa a las políticas editoriales y derechos de edición, contabilizándose un 52% que ofrecen información sobre SHERPA/RoMEO y 49% sobre DULCINEA. No obstante, no se ha identificado ninguna integración de estos servicios de información en los flujos de archivo existentes, con las salvedades apuntadas por la metodología de observación empleada.

### Gestores de Referencias bibliográficas

La interoperabilidad de estos sistemas con los gestores de referencias bibliográficas está moderadamente extendida, existiendo además repositorios con un alto número de integraciones con diversos gestores, ofreciendo múltiples opciones de exportación-integración a los usuarios.

Repositorios con Integración	23	19	9	2	2	2
Nº de Gestores de Referencias soportados	0	1	2	3	5	6

Tabla 3: Nº de Gestores Bibliográficos soportados

BibTex	ProCite	Reference Manager	RIS	EndNote	Mendeley	RefWorks
3	4	4	5	8	12	29

Tabla 4: Nº de repositorios integrando un determinado Gestor de Referencias

### Sistemas de Estadísticas de uso del repositorio

La estadísticas identificadas en los repositorios, ofrecidas a los usuarios finales, corresponden a los sistemas internos propios o extensiones de los mismos, no habiendo encontrado un ejemplo de integración con sistemas de estadísticas externas. No obstante señalar que 33 repositorios (58%) incorporan el código de Google Analytics, por lo que se puede suponer el análisis de los datos de tráfico por los responsables de estos repositorios.

### Sistemas de métricas de impacto y alométricas

Encontramos diez repositorios (18%) ofreciendo servicios integrados de métricas de impacto y alométricas, siendo Cited-by Scopus el más habitual, y Altmetrics y WOS menos comunes.

Repositorios con Integración	47	5	4	1
Nº de Sistemas soportados	0	1	2	3

Tabla 5: Nº de sistemas de métricas y alométricas soportados

WOS	Altmetrics	Cited-By Scopus
3	5	8

Tabla 6: Nº de repositorios integrando un determinado sistema

## CONCLUSIONES

El resultado final, además de caracterizar el estado del arte de la integración de los repositorios digitales, apunta potencialidades de evolución para que los responsables de estos sistemas puedan dibujar el mapa de alternativas posibles para convertir los repositorios en puntos centralizados para el depósito, difusión y preservación del material académico y científico de las instituciones universitarias.

Cada vez hay más posibilidades tecnológicas de integración de los repositorios en el complejo ecosistema de ayuda y soporte a los investigadores. Los repositorios estudiados presentan avances generales en integraciones débiles, claramente con recolectores, pero también con sistemas de gestión de referencias bibliográficas y en menor medida con sistemas de métricas de impacto y altmetrías.

Consideramos que se necesita implementar e integrar fuertemente los repositorios en el resto de sistemas soporte del trabajo de la comunidad académica, demostrando el valor de los repositorios como infraestructura perteneciente y conectada al resto de la infraestructura de investigación.

Para que los repositorios sean herramientas de alto uso, relevantes a usuarios, necesitan implementar de forma continua los servicios de valor que éstos demandan. Dicho de otra forma, los repositorios que no proporcionen los complejos servicios interconectados que se requieren, podrán cumplir con su misión original de acceso, difusión y preservación de la producción científica, pero posiblemente tendrán poco impacto y relevancia en el conjunto del sistema investigador.

## BIBLIOGRAFÍA

Allinson, J., Francois, S., & Lewis, S. (2008). SWORD: Simple Web-service Offering Repository Deposit. *Ariadne Issue 54*.

Altmetric. (s.f.). <http://api.altmetric.com/>. Obtenido de <http://api.altmetric.com/>

AWStats. (s. f.). AWStats - Free log file analyzer for advanced statistics. Recuperado el 28 de febrero de 2015, a partir de <http://www.awstats.org/>

Barrueco, J., & Subirats Coll, I. (2003). Open archives initiative. Protocol for metadata harvesting (OAI-PMH): descripción, funciones y aplicaciones de un protocolo. *El Profesional de la Información*, 12(2), 99-106.

Bell, S., & Sarr, N. (2010). Case Study: Re-Engineering an Institutional Repository to Engage Users. *New Review of Academic Librarianship*, 16(1), 77-89.

BioMed Central Ltd. (s.f.). <http://www.biomedcentral.com/>. Obtenido de <http://www.biomedcentral.com/>

Borrego, Á. (2013). *Sistemes d'identificació unívoca d'investigadors*. Universitat de Barcelona, Barcelona.

Borrego, Á. (2014). Altmétricas para la evaluación de la investigación y el análisis de necesidades de información. *El Profesional de la Información*, 23(4), 352-358.

- Confederation of Open Access Repositories, COAR. (2012). <https://www.coar-repositories.org>. Obtenido de <https://www.coar-repositories.org/files/COAR-Current-State-of-Open-Access-Repository-Interoperability-26-10-2012.pdf>
- Confederation of Open Access Repositories, COAR. (2013). *Incentivos, integración y mediación: prácticas sostenibles para poblar repositorios*. COAR.
- Confederation of Open Access Repositories, COAR. (2015). <https://www.coar-repositories.org>. Obtenido de [https://www.coar-repositories.org/files/Roadmap\\_final\\_formatted\\_20150203.pdf](https://www.coar-repositories.org/files/Roadmap_final_formatted_20150203.pdf)
- Cornell University Library. (s.f.). <http://arxiv.org>. Obtenido de <http://arxiv.org/help/api/index>
- Cybermetrics Lab; Consejo Superior de Investigaciones Científicas, CSIC. (s.f.). <http://repositories.webometrics.info>. Obtenido de <http://repositories.webometrics.info/en/Europe/Spain>
- DULCINEA. (s.f.). *DULCINEA*. Obtenido de <http://www.accesoabierto.net/dulcinea/>
- DURASPACE. (s. f.). DSpace Google Analytics Statistics - DSpace 5.x Documentation - DuraSpace Wiki. Recuperado 28 de febrero de 2015, a partir de <https://wiki.duraspace.org/display/DSDOC5x/DSpace+Google+Analytics+Statistics>
- Elsevier. (s.f. -a). *Elsevier developers*. Obtenido de [http://dev.elsevier.com/tecdoc\\_ir\\_cris\\_vivo.html](http://dev.elsevier.com/tecdoc_ir_cris_vivo.html)
- Elsevier. (s.f. -b). <http://dev.elsevier.com>. (Elsevier, Editor) Obtenido de [http://dev.elsevier.com/cited\\_by\\_scopus.html](http://dev.elsevier.com/cited_by_scopus.html)
- FECYT. (s. f.). Repositorios Recolectados. Recolecta. Recuperado el 28 de febrero de 2015, a partir de <http://recolecta.fecyt.es/repositorios-recolectados>
- Flowerfire, Inc. (s. f.). Sawmill - Universal log file analysis and reporting. Recuperado el 28 de febrero de 2015, a partir de <https://www.sawmill.net/>
- Gordon, D. (2003). What is Interoperability and why does it matter? *MapInfo Magazine*, 7(4), 20.
- Lewis, S., De Castro, P., & Jones, R. (2012). SWORD: Facilitating Deposit Scenarios. *D-Lib Magazine*, 18(1/2).
- Open Archives Initiative. (s.f.). *Open Archives Initiative Protocol - Object Exchange and Reuse*. Obtenido de <http://www.openarchives.org/ore/>
- ORCID. (s.f.). *ORCID. Connecting Research and Researchers*. Obtenido de <http://support.orcid.org/knowledgebase>
- Plum Analytics, Inc. (s.f.). <https://plu.mx/developers>. Obtenido de <https://plu.mx/developers>
- Red de Bibliotecas Universitarias, REBIUN. (s.f.). <http://www.rebiun.org/repositorios>. Obtenido de <http://www.rebiun.org/repositorios/Paginas/Directorio-de-Repositorios-Institucionales-REBIUN.aspx>
- Serrano Vicente, R., Melero Melero, R., & Abadal, E. (2014). Indicadores para la evaluación de repositorios institucionales de acceso abierto. *Anales de Documentación*, 17(2), 12.

Sherpa/Romeo. (s.f. -a). <http://www.sherpa.ac.uk>. Publisher copyright policies & self-archiving. Obtenido de <http://www.sherpa.ac.uk/romeo/>

Sherpa/Romeo. (s.f. -b). <http://www.sherpa.ac.uk>. SHERPA/RoMEO Application Programmers' Interface. Obtenido de <http://www.sherpa.ac.uk/romeo/api.html>

Thibodeau, K. (2007). If you build it, will it fly? Criteria for success in a digital repository. *Journal of Digital Information*, 8(2), 7.

Van de Sompel, H., L. Nelson, M., Lagoze, C., & Warner, S. (2004). Resource Harvesting within the OAI-PMH Framework. *D-Lib Magazine*, 10(12).

Westell, M. (2006). Institutional repositories: proposed indicators of success. *Library Hi Tech*, 24(2), 211-226.

Zenodo. (s.f.). <http://zenodo.org>. Obtenido de <http://zenodo.org/dev>