

Servicios para repositorios en el Reino Unido: el proyecto **UK RepositoryNet+**

Pablo De-Castro

23 septiembre 2013

De-Castro, Pablo (2014). "Servicios para repositorios en el Reino Unido: el proyecto *UK RepositoryNet+*". *Anuario ThinkEPI*, v. 8, pp. 215-219.



Resumen: El proyecto *UK RepositoryNet+* (también conocido como *RepNet*), financiado por el *Joint Information Systems Committee (JISC)* en el Reino Unido, pretende implementar una infraestructura socio-técnica para mejorar la red de repositorios institucionales del país. La primera fase del mismo -fase de proyecto- transcurrió desde agosto de 2011 hasta julio de 2013, y la iniciativa ha sido prolongada durante dos años más (aunque su nombre cambiará a lo largo de esta segunda fase) durante los cuales se propone convertirse en un servicio. Hasta el momento *RepNet* ha agrupado bajo una plataforma única diversos proyectos y/o servicios para repositorios previamente existentes, y ha elaborado una metodología basada en ITIL para transformar proyectos en servicios. Se describen las diferentes áreas de trabajo que ha identificado el proyecto, los servicios que cubre y el trabajo pendiente.

Palabras clave: Repositorios institucionales, Acceso abierto, Servicios para repositorios, Metadatos, Gestión de la información científica.

Title: *An attempt to provide new services to the repository network in the UK: the UK RepositoryNet+ Project*

Abstract: The JISC-funded *UK RepositoryNet+* (aka *RepNet*) project aims to build a socio-technical infrastructure for enhancing the institutional repository network at UK Higher Education institutions. The *RepNet* ran from Aug'2011 to Jul'2013 in its project stage, and has now been extended two years (undergoing a name change along the way), during which it aims to become a new service. So far, the *RepNet* has brought together existing repository projects and/or services under a common umbrella and has devised an ITIL-based workflow for actually turning repository projects into new repository services. This note describes the different repository working areas the *RepNet* has identified, the services it covers and the pending work.

Keywords: Institutional Repositories, Open Access, Repository Services, Metadata, Research Information Management.

1. Introducción

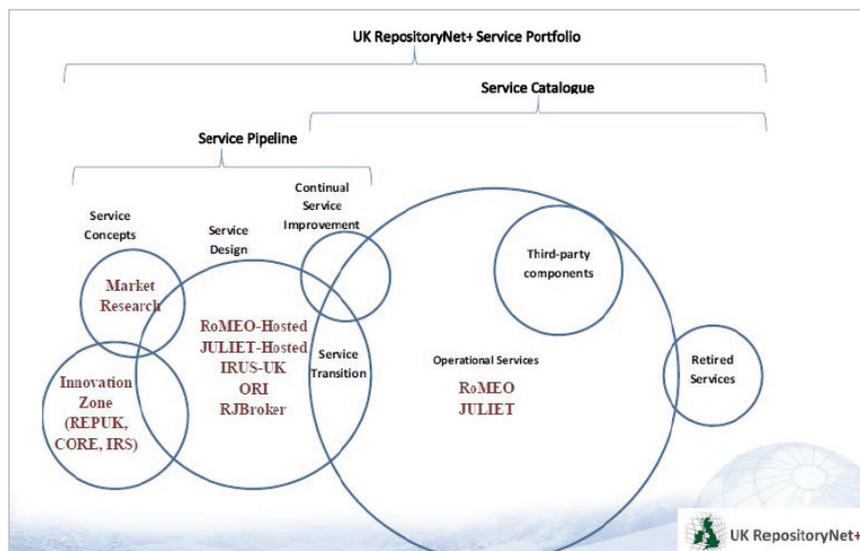
Después de la ya tardía *Declaración de la Alhambra* (mayo 2010), continúan llegando desde España noticias sobre nuevas declaraciones en apoyo del acceso abierto a nivel institucional. Aunque no completamente desprovistas de utilidad –especialmente si redundan en una mejor dotación de medios técnicos y humanos para los equipos que tratan de implantar los objetivos citados en dichos textos–, estas declaraciones carecen de sentido si se limitan a ser meras expresiones de apoyo a una iniciativa próxima a cumplir diez años desde su lanzamiento¹.

Una vez que, como fruto del trabajo de muchos profesionales en las bibliotecas universitarias y de centros de investigación de todo el mundo, se ha alcanzado un grado de consolidación de la red de repositorios de acceso abierto que no admite vuelta atrás, el siguiente paso es aventu-

rarse en el desarrollo de servicios sobre esa capa de infraestructura que atiendan a las necesidades de académicos e investigadores y de sus instituciones.

Este es el espíritu que ha guiado el proyecto *UK RepositoryNet+* en el Reino Unido, que se autodefine como "una iniciativa para la creación de una infraestructura socio-técnica que soporte el depósito, la curación y la difusión en acceso

"UK RepositoryNet+ es una iniciativa para la creación de una infraestructura socio-técnica que soporte el depósito, la curación y la difusión en acceso abierto de la literatura de investigación"



Uk RepositoryNet+ Service Portfolio

http://repositorynet.ac.uk/sites/default/files/documents/images/RepNet_and_ITIL.pdf

abierto de la literatura de investigación”.
<http://repositorynet.ac.uk>

Mucho se ha hablado en este último año de la “errónea apuesta del Gobierno británico por un modelo insostenible de acceso abierto “dorado” (*gold open access*) financiado mediante cuotas por procesamiento de artículos detraídas de los magros presupuestos disponibles para la investigación”. Sin pretender que dicha afirmación sea completamente errónea, es preciso tener también en cuenta la cuantiosa inversión (con cifras de siete dígitos en libras esterlinas) realizada simultáneamente en una investigación sobre las vías de consolidación de la ruta verde y los repositorios de acceso abierto sin parangón en Europa² a través de *RepNet*. Por el contrario, este proyecto apenas aparece mencionado en las acaloradas discusiones *gold vs green* que vienen teniendo lugar desde hace algún tiempo en las listas de distribución de la disciplina.

“RepNet propone la construcción de un agregador de contenidos de toda la red de repositorios del Reino Unido”

Esto se debe principalmente al hecho de que, frente a la simplicidad de una política de acceso abierto concreta que es fácil juzgar y aprobar o condenar, el análisis de un proyecto tan complejo como *RepNet* requiere un conocimiento profundo de los retos técnicos que plantean los diferentes servicios para repositorios y de los enfoques adoptados para resolverlos por los equipos encargados de su desarrollo. De esta manera, aunque

prácticamente ausente de las -frecuentemente bizantinas- discusiones entre los abogados del acceso abierto, *RepNet* ha sido muy comentado y debatido por la comunidad de *repository managers* en el Reino Unido, que es la encargada de implantar las a menudo cambiantes, cuando no contradictorias, políticas emanadas desde las distintas instancias administrativas a nivel institucional, regional o nacional.

Tal como se presenta en la página principal del proyecto, el desarrollo de servicios sobre la capa de repositorios se sustenta sobre un análisis previo de las necesidades de los diferentes actores implicados (instituciones, agencias de financiación, investigadores...)

y sobre la definición de una serie de áreas de trabajo en las cuales es perentorio proporcionar nuevas funciones para garantizar la continuidad de los repositorios de acceso abierto. Todo ello sucede en un momento en el que las exigencias para cumplir con los requisitos de aportación de información científica que plantea el *Research Excellence Framework* (REF) –el ejercicio de evaluación científica que se llevará a cabo en el Reino Unido en 2014– hacen que muchas instituciones hayan optado por implantar sistemas CRIS, que a menudo amenazan con reemplazar a los repositorios de acceso abierto, pese a basarse en un enfoque mucho más centrado en la gestión de información científica que en el acceso abierto como tal³.

2. Áreas de actividad de RepNet

Las áreas de actividad de *RepNet* a nivel de identificación, diseño, desarrollo e implantación de servicios para repositorios son las siguientes:

Agregación de contenidos

RepNet propone la construcción de un agregador de contenidos de toda la red de repositorios del país. A diferencia de muchos otros países en los que esta función existe desde hace tiempo, en el Reino Unido no se ha consolidado ninguna de las diferentes iniciativas que han desarrollado prototipos para la agregación de contenidos.

Esta desventaja a nivel de infraestructura tiene la contrapartida de que una plataforma construida en este momento puede ofrecer funciones mucho más avanzadas que las que poseen las plataformas desarrolladas con anterioridad. Algunas de estas son la minería de datos sobre los textos completos de los documentos archivados con asignación au-

tomática de descriptores, la detección e integración de duplicados a partir de una estrategia similar de análisis del texto completo de los contenidos y la detección de registros *metadata-only* (sin texto completo asociado), incluso aunque contengan un archivo pdf por defecto o *default dummy file* para indicar que el texto completo no está disponible.

Teniendo en cuenta que la adopción de las directrices *Driver (Digital Repository Infrastructure Vision for European Research)* ha sido muy escasa en el Reino Unido (lo que ha llevado a su vez a niveles de cumplimiento inusualmente bajos de los estándares de *OpenAI-RE*), una agregación puede ofrecer una novedosa función de validación de esquemas de metadatos, aplicando criterios muy avanzados como los de detección de las versiones de los artículos archivados o la agregación de información de financiación de los trabajos.

Generación de informes y comparativa de plataformas

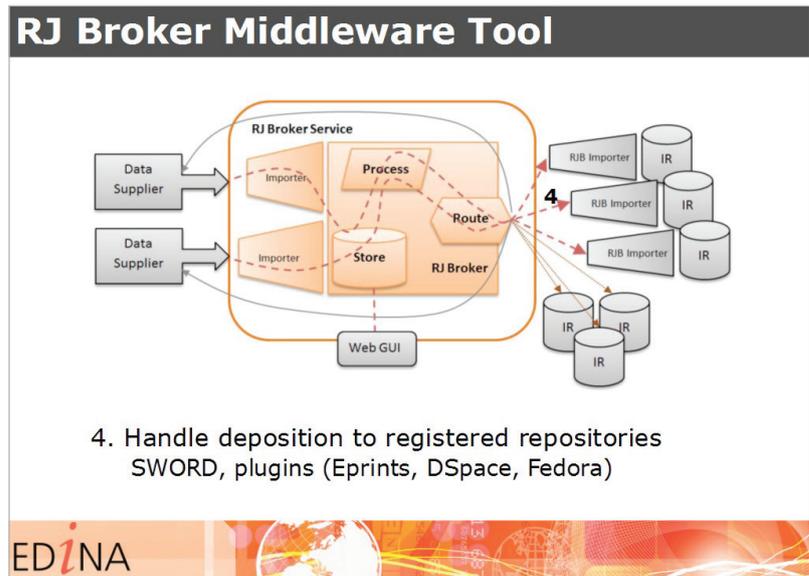
En el área de *reporting*, *RepNet* viene operando el proyecto *IRUS-UK*, siguiendo un modelo común de incubación de servicios externalizados de acuerdo con la metodología *ITIL*⁴.
<http://irus.mimas.ac.uk>

IRUS-UK es un proyecto desarrollado en el *Centro de Datos Mimas* de la *Manchester University* para recolectar estadísticas de uso de múltiples repositorios armonizadas de acuerdo con el estándar *Counter*. A mediados de septiembre de 2013, *IRUS-UK* recoge datos de 40 repositorios institucionales –aproximadamente un tercio de la red nacional– y continúa extendiendo su cobertura, limitada por el momento a *EPrints* (29) y *DSpace* (11), en tanto el equipo de desarrollo trabaja en el módulo de intercambio de datos para *Fedora* y otras plataformas.

Además de permitir la comparación para diferentes plataformas y tipos de documentos, el objetivo de *IRUS-UK* es obtener una estimación de las estadísticas de uso agregadas para toda la red, en la confianza de que los niveles de uso globales resultarán un argumento convincente para garantizar la utilización continuada de la misma por parte de autores e instituciones.

Depósito automático de contenidos

El proyecto *Repository Junction Broker (RJB)* es una iniciativa para la transferencia automatizada de contenidos a la red de repositorios a través del protocolo *Sword*. Después de varios años de



4. Handle deposition to registered repositories SWORD, plugins (Eprints, DSpace, Fedora)

http://edina.ac.uk/projects/docs/RJBroker_Final_Report_v1.1.pdf

trabajo, el proyecto *RJB* se incluyó como parte de los servicios a prestar por parte de *RepNet*, y ha sido bajo este paraguas cuando ha comenzado a funcionar como servicio en fase piloto desde mediados de este año.

RJB pretende consolidar una base de proveedores de contenido, fundamentalmente a nivel de artículos de revista, que puedan ser distribuidos bien como registros sólo de metadatos o como metadatos más texto completo, a los diversos repositorios institucionales correspondientes a las afiliaciones de los autores de cada artículo concreto.

“Es preciso mejorar la visibilidad de los contenidos de los repositorios en la Red”

En un principio el *RJB* ha firmado acuerdos con el repositorio temático *EuropePMC* y con *Nature Publishing Group* para distribuir los contenidos de ambos proveedores como proyecto piloto (el primero según el modelo *metadata-only* y el segundo transfiriendo *metadata* más *full-text*, lo que requiere el compromiso por parte de los repositorios receptores de no difundir los textos completos antes de la fecha de embargo).

Un aspecto clave de la operación de este servicio es su naturaleza internacional por defecto: dado que los autores de los artículos son con frecuencia internacionales, basta con que los repositorios institucionales susceptibles de recibir información estén registrados en el servicio para que automáticamente puedan recibir los contenidos (previa instalación de *Sword*) con independencia del país en el que estén ubicados.

Enriquecimiento de metadatos

El área de *metadata enhancement* es posiblemente la más amplia de las que aborda el proyecto UKRepNet. Fruto de las investigaciones previas sobre necesidades de los diferentes ámbitos implicados, se puso de manifiesto la existencia de estrategias para la asignación de metadatos puestas en práctica por repositorios aislados (por ejemplo en el ámbito de la preservación de contenidos), que no se difundían al resto de la red.

Vista la necesidad de armonizar el desarrollo de toda la red al mismo ritmo, se puso en marcha la iniciativa *Rioxx* para el desarrollo e implantación de un *application profile* que permitiera la incorporación conjunta de metadatos sobre financiación (algo que ya abordaba *OpenAIRE* para los proyectos FP7), sobre aspectos específicos relativos al acceso y sobre identificadores como *Orcid*. <http://rioxx.net>

Las iniciativas preliminares para la incorporación de estos metadatos avanzados a los repositorios han comenzado a difundirse recientemente (**Nixon; Ashworth; McCutcheon, 2013**), de modo que puedan gradualmente adoptarse de manera conjunta por parte de toda la red.

Registro de repositorios

Los dos principales directorios de repositorios existentes en la actualidad, *OpenDOAR* y *Roar*, mantenidos respectivamente por las universidades de *Nottingham* y *Southampton*, proporcionan una cobertura más que aceptable de la red mundial de repositorios.

Sin embargo, ninguno de ellos proporciona una cobertura completa de la red. Por este motivo y también para actualizar el perfil que los directorios proporcionan sobre las plataformas que indexan, se ha puesto en marcha como parte de UKRepNet el proyecto *Open Access Repository Registry (OARR)*¹⁰.

<http://bit.ly/LeXGjp>

Este proyecto pretende actualizar la información de *OpenDOAR* cubriendo en mayor detalle las características de los repositorios, en un momento en que tanto la implantación generalizada de sistemas CRIS como el creciente número de repositorios de datos de investigación están introduciendo cambios significativos en el sector.

El nuevo directorio, cuyo proyecto lidera el equipo *CRC-Sherpa* en la *Universidad de Nottingham*, se alojará eventualmente en los servidores de *RepNet* junto a otros servicios proporcionados por *Sherpa* tales como *RoMEO*, *Juliet* o más recientemente, *Fact*. De hecho, una de las líneas para el diseño de nuevos servicios para repositorios pasa por explotar las sinergias entre estas aplicaciones gestionadas de manera integrada.

Localización de la información

Una de las cuestiones más problemáticas de los repositorios hace referencia a su escasa visibilidad en la red. Junto a la creación de esquemas de metadatos suficientemente comprensivos que puedan servir los propósitos de la *discoverability*, la línea de trabajo orientada a la mejora de la visibilidad de los contenidos pretende sobre todo optimizar los ratios de indexación de los materiales archivados en la red de repositorios del Reino Unido por parte de motores de búsqueda como *Google scholar* o *Microsoft academic search*.

Sea a través de la identificación de buenas prácticas a nivel de repositorio individual o bien a través de la indexación masiva de una agregación de contenidos, es preciso mejorar la visibilidad de los contenidos de los repositorios en la Red, así como identificar su procedencia de modo que el usuario final de la información pueda conocer y valorar la labor realizada desde estas plataformas.

Sea a través de la identificación de buenas prácticas a nivel de repositorio individual o bien a través de la indexación masiva de una agregación de contenidos, es preciso mejorar la visibilidad de los contenidos de los repositorios en la Red, así como identificar su procedencia de modo que el usuario final de la información pueda conocer y valorar la labor realizada desde estas plataformas.

Preservación/continuidad de acceso

Sin entrar directamente en el área de la preservación digital, cubierta por otros progra-

<http://repositorynet.ac.uk/>

mas y proyectos del Jisc, como Spruce⁵, el proyecto UKRepNet sí se planteó en cambio ofrecer alguna clase de servicio para la red de repositorios en el sentido de asegurar la continuidad de acceso a los contenidos archivados en la misma.

Para ello UKRepNet trabaja sobre la extensión a los materiales archivados en acceso abierto del modelo Lockss, ya empleado con éxito para la gestión de la continuidad en el acceso a materiales obtenidos a través de suscripción por parte de las bibliotecas (Rusbridge, 2012). Este modelo se basa en el archivo periódico de los contenidos en una red de servidores distribuidos (las Lockss boxes) gestionada por las instituciones.

3. Servicios de nueva creación

Además del énfasis en la integración y ulterior desarrollo de los servicios para repositorios ya existentes, el proyecto UKRepNet pretende también abordar el diseño, desarrollo e implantación de una serie de nuevos servicios. Para ello adopta el modelo para la construcción de una infraestructura (de servicios) basada en datos o *data-driven infrastructure* (Hammond, 2013), que permita plantear la puesta en marcha de servicios de nueva creación largamente demandados por la comunidad, tales como herramientas para la monitorización del cumplimiento de mandatos de acceso abierto.

La creación de nuevos servicios se lleva a cabo mediante el establecimiento de *partnerships* con instituciones concretas que permitan el ensayo y testeo de desarrollos piloto. Así, la iniciativa Stars (De-Castro; Proven, 2013), llevada a cabo en colaboración con la Universidad de St. Andrews y el Scottish Digital Library Consortium (SDLC) se ha planteado como una prueba piloto para la implantación del conjunto de servicios que una plataforma como UKRepNet puede ofrecer a una institución y un repositorio específicos.

4. Notas

1. La *Declaración de Berlín*, publicada por la *Sociedad Max Planck* en octubre de 2003, puede considerarse como el pistoletazo de salida del movimiento del acceso abierto con la opción que ofrecía a organismos académicos y de investigación para suscribirla de manera institucional. De hecho, la *Semana del acceso abierto* se celebra anualmente en el mes de octubre como conmemoración de la publicación de esta *Declaración*.

2. Sólo el proyecto europeo OpenAIRE plantea un nivel de objetivos de similar amplitud y ambición a los de RepNet a nivel de servicios a desarrollar sobre la red de



<http://www.openaire.eu/>

repositorios de acceso abierto existente en la actualidad.
<http://www.openaire.eu>

3. En relación con el impacto sobre las instituciones del ejercicio de recopilación de información científica para el REF2014, véase la excelente presentación "I am turning enterprise" realizada por **Chris Keene** (*repository manager* en la *University of Sussex*), en la reciente conferencia "Repository Fringe 2013", celebrada en Edimburgo el mes de agosto de 2013.
<http://bit.ly/159PtW9>

4. Ver referencia a ITIL en la sección de preguntas frecuentes de RepNet.
<http://www.repositorynet.ac.uk/?q=content/faq>

5. *Sustainable Preservation Using Community Engagement* (Spruce).
<http://bit.ly/1aXY1Vd>

El texto de esta nota ThinkEPI está disponible con hipervínculos y gráficos en *GrandIR Blog*:
<http://bit.ly/1aXa95w>

5. Bibliografía

De-Castro Pablo; Proven, Jackie (2013). "The STARS shared initiative: delivering repository services in an advanced CRIS/IR environment". *RepositoryFringe* 2013.
<http://slidesha.re/1eXH2nl>

Hammond, Max (2012). *Preparing for data-driven infrastructure*. JISC Observatory TechWatch report, septiembre.
<http://bit.ly/1a8SuXe>

Nixon, William J.; Ashworth, Susan; McCutcheon, Valerie (2013). "Enlighten: Research and APC funding workflows at the University of Glasgow". *Insights: the UKSG journal*, v. 26, n. 2, pp. 159-167.
<http://dx.doi.org/10.1629/2048-7754.80>
<http://eprints.gla.ac.uk/183882>

Rusbridge, Adam (2013). "UK LOCKSS alliance case studies now available". *UK LOCKSS Alliance*, 5 abril.
<http://bit.ly/18Gkw9j>

UK RepositoryNet+ (2013). "RJ Broker delivers its first test transfers". *Developing UK RepositoryNet+*, 27 marzo.
<http://bit.ly/16dsJmq>