

XII Jornadas de Gestión de la Información.
Valor económico de la información: mercados, servicios y rentabilidad.
SEDIC, 18-19 de noviembre de 2010

REPEC COMO MODELO DE SOSTENIBILIDAD DE UNA BIBLIOTECA ABIERTA DE CONTENIDOS CIENTÍFICOS

REPEC AS A SOSTENIBILITY MODEL OF AN OPEN LIBRARY OF SCIENTIFIC CONTENTS

Barrueco Cruz, José Manuel. Servei de Biblioteques, Universitat de València. C/ Universidad, 2 46003 VALENCIA, barrueco@uv.es; Krichel, Thomas. Long Island University, Palmer School of Library and Information Science y Novosibirsk State University, Faculty of Information Technology, 720 Northern Boulevard Brookville NY 11548–1300 United States of America, krichel@openlib.org.

Resumen: RePEc (Research Papers in Economics) es la mayor biblioteca digital en acceso abierto que existe actualmente. Contiene más de 900.000 documentos en formato electrónico y el volumen de descargas de artículos alcanza una media de 700.000 al mes durante el último año. En sus trece años de existencia se ha convertido en una fuente de información indispensable para acceder a los últimos resultados de investigación en Economía. La arquitectura de RePEc es de carácter distribuido con más de 1200 instituciones que colaboran aportando datos para formar un dataset de acceso público sobre el que se desarrolla una serie de servicios de valor añadido como: alertas de novedades, análisis de citas, etc. Todos los contenidos son reutilizables pudiendo ser aplicados por cualquier comunidad que quiera desarrollar nuevos servicios. El modelo de negocios de RePEc está asentado sobre los mismos principios que mueven la comunidad de software libre: colaboración voluntaria a través de unos procedimientos comunes para aportar metadatos a un dataset común, de libre acceso y reutilizable. En un momento en que otros modelos de negocios, de iniciativas como por ejemplo arXiv, han mostrado problemas de sostenibilidad por el alto coste que supone su mantenimiento, es interesante estudiar el modelo de RePEc porque se ha mostrado sólido y sostenible al basarse en una comunidad de usuarios con unos intereses compartidos. En esta comunicación hacemos una introducción a RePEc, analizamos su modelo de negocios y desarrollamos un análisis de las fortalezas y debilidades que esta iniciativa tiene planteadas.

Abstract:

RePEc (Research Papers in Economics) is the largest academic open access digital library world wide. It references more than 900.000 documents in electronic format. Download of articles and papers it is around 700.000 a month in the last year. In its thirteen years of life, it has become in an indispensable information source of the latest research results in Economics. The RePEc architecture is distributed. More than 1200 institutions collaborate to build a public dataset. Several value added services use the data: alerts about new contents, citation indexing, etc. All contents are re-usable so that any institution may use them for its own community or to build new services for end users. The RePEc business model is based on the same principles as the open source movement: volunteer collaboration through common procedures to add metadata to a shared, open access and re-usable dataset. At a time when the business models of other similar initiatives like arXiv are under review due to sustainability problems, it is interesting analyze the RePEc model. This model has proven strong and sustainable since it is based in a community of users with shared interests. In this paper we introduce RePEc, analyze its business model study its strengths and weaknesses.

Palabras clave: Acceso abierto; repositorios; modelos de negocios; código abierto

Keywords: Open access; repositories; business models; open source

1.- Introducción

El movimiento de acceso abierto es un tema candente en el ámbito de las bibliotecas universitarias. Se entiende por tal, según la declaración de Budapest¹, “... la libre disponibilidad de la literatura científica en Internet, permitiendo a cualquier usuario su lectura, descarga, copia, impresión, distribución o cualquier otro uso legal de la misma, sin ninguna barrera financiera, técnica o de otro tipo”. Es decir, se trata de conseguir que cualquier investigador tenga acceso a los documentos que necesita para su trabajo independientemente de donde se encuentre o de cuáles sean los recursos económicos de la institución donde trabaja. La misma declaración establece dos vías para alcanzar el acceso abierto: la vía verde, consistente en persuadir a los científicos e investigadores de que, además de publicar en las revistas tradicionales, almacenen una copia de sus trabajos en repositorios institucionales o temáticos; y la vía oro, que añade a este presupuesto la intención de crear o convertir las revistas ya existentes en revistas de acceso abierto.

Para que la vía verde sea una realidad es necesario que se desarrolle una infraestructura de información que permita a los investigadores almacenar y localizar sus trabajos una vez que han sido publicados. Tal infraestructura se describe en la Open Archives Initiative² y se asienta en dos pilares básicos: los almacenes de documentos científicos o repositorios y los recolectores o proveedores de servicios a los usuarios finales basados en los primeros.

Existen dos tipos de repositorios, los institucionales destinados a almacenar la producción científica de una institución (Universidad, Departamento, Instituto, etc.) y los temáticos, destinados a albergar la producción generada a nivel internacional en una disciplina concreta. Las bibliotecas se han lanzado de forma entusiasta al desarrollo de los primeros. En cinco años se han abierto unos 25 repositorios institucionales en nuestro país, mientras que a nivel internacional el directorio OpenDOAR³ recoge más de 1600. Frente a los institucionales, el número de repositorios temáticos es muy reducido, pero con un elevado volumen de contenidos y con una trayectoria consolidada pues en su mayor parte nacieron durante la década de los noventa. Entre ellos podemos citar: arXiv⁴ (Física), CiteSeer⁵ (Informática), RePEc⁶ (Economía), E-LIS⁷ (Documentación).

Por su parte, los proveedores de servicios deben ser capaces de integrar los contenidos de los repositorios en un sistema centralizado que provea a los investigadores de servicios de valor añadido tales como motores de búsqueda, categorización de contenidos, alertas, índices de citas, etc. Los proveedores de servicios también son pocos: OAIster⁸ (OCLC), BASE⁹ (Bielefeld Academia Search Engine, Universidad de Bielefeld), ScientificCommons¹⁰ o Recolecta¹¹ (FECYT, REBIUN) son algunos ejemplos.

Todas estas iniciativas nacen con el objetivo de poner los contenidos científicos en libre acceso, sin ninguna restricción financiera para el usuario final, lo cual implica que los servicios ofrecidos deban ser gratuitos para el mismo. Este principio fundamental plantea el reto de buscar vías de financiación para los proyectos de acceso abierto que aseguren su sostenibilidad sin recurrir al sistema de suscripciones o de pago por uso.

Aunque la literatura producida en los últimos años es considerable, faltan estudios rigurosos que aborden el tema de la financiación y sostenibilidad de las iniciativas de acceso abierto. Las formas de financiación de los repositorios institucionales han sido los recursos propios (Henty, 2007), aportados por las instituciones de quienes dependen, y las subvenciones procedentes de organismos nacionales o internacionales tales como la Unión Europea. Los repositorios temáticos y los proveedores de servicios siguen un esquema similar de financiación. En muchos

1 <http://www.soros.org/openaccess>
2 <http://www.openarchives.org/>
3 <http://www.opendoar.org/>
4 <http://arxiv.org/>
5 <http://citeseerx.ist.psu.edu/>
6 <http://repec.org>
7 <http://eprints.rclis.org>
8 <http://www.oclc.org/oaister/>
9 <http://base.ub.uni-bielefeld.de/en/index.php>
10 <http://en.scientificcommons.org/>
11 <http://recolecta.net/buscador/>

casos surgieron como un proyecto de investigación con una financiación a corto término y a medida que fueron creciendo se vieron en la necesidad de buscar recursos alternativos.

Si bien la viabilidad de los primeros está asegurada a corto plazo ya que el repositorio se considera una herramienta útil para mejorar la visibilidad de la universidad, la de los segundos está cuestionada por la necesidad de buscar recursos externos para seguir proporcionando un servicio gratuito.

En el presente trabajo hacemos un estudio de la viabilidad de un caso concreto de repositorio temático especializado en Economía: RePEc (Research Papers in Economics), creado por los autores de esta comunicación en 1997. Nuestro objetivo es analizar el modelo de negocios de RePEc poniéndolo en relación con otras iniciativas similares. Se intentará demostrar que un modelo basado en la colaboración y la distribución de costes puede ser viable para proporcionar una serie de servicios gratuitos a la comunidad académica. Además se ofrece un análisis de las fortalezas y debilidades que esta iniciativa tiene planteadas.

El resto del trabajo se organiza de la siguiente forma. La sección dos analiza las características que hacen de RePEc un caso particular de repositorio temático. La sección tres analiza modelos de negocios en el ámbito de los principales proveedores de servicios y archivos temáticos. La cuatro se centra en la descripción del modelo de negocios de RePEc. La sección cinco desarrolla un análisis de sus fortalezas y debilidades. Finalmente, la seis concluye el trabajo.

2.- RePEc, un repositorio poco común

RePEc es la mayor biblioteca digital en acceso abierto que existe actualmente. En Agosto de 2010 contiene metadatos sobre más de 900.000 documentos de investigación especializados en Economía, fundamentalmente artículos y working papers. Unos 800.000 de ellos están disponibles en formato electrónico, aunque no todos en acceso abierto. Entre ellos se encuentran las principales revistas internacionales de la disciplina y documentos de trabajo producidos por departamentos universitarios y organismos tales como OCDE, FMI o Banco Mundial.

En sus trece años de existencia se ha convertido en una fuente de información indispensable para acceder a los últimos resultados de investigación en Economía. Esta afirmación la avala el hecho de que el volumen de descargas de documentos ha alcanzado una media de 700.000 al mes durante el último año según se refleja en las estadísticas de uso¹².

Entre las características que hacen de RePEc una iniciativa particular podemos citar:

- Su arquitectura distribuida. Está formada por más de 1200 instituciones (universidades, organismos oficiales, editoriales, etc.) que colaboran aportando metadatos para formar un *dataset* de acceso público sobre el que se desarrolla una serie de servicios de valor añadido como: búsquedas, alertas de novedades, análisis de citas, etc. En este sentido RePEc es un antecedente de la arquitectura definida por la Open Archives Initiative. El principio fundamental de RePEc es: muchos archivos o repositorios forman un conjunto de metadatos que se comunican a unos agregadores o proveedores de servicios los cuales los utilizarán para proporcionar servicios de valor añadido a la comunidad académica.
- RePEc no es un repositorio. A pesar de que aparece en el puesto tercero del “Ranking web de repositorios del mundo” elaborado por el Laboratorio de Cibermetría del CSIC¹³, no podría considerarse como tal, puesto que no es un almacén de documentos sino un conjunto de metadatos. No existe un depósito central donde se guarden los ficheros que contienen los textos de los documentos, sino que estos están distribuidos en los servidores de las instituciones colaboradoras, quienes en todo momento mantienen el control sobre los mismos.

¹² <http://logec.repec.org>

¹³ http://repositories.webometrics.info/index_es.html

- RePEc es una biblioteca abierta (Krichel, 2000) basada en el principio de la reutilización de contenidos. Todos los datos y metadatos son reutilizables por cualquiera, bajo determinadas condiciones: (1) no pueden ser incluidos en servicios de pago ni se puede cobrar por el acceso a los mismos, (2) siempre que se utilicen debe proporcionarse un número mínimo por cada ítem, (3) a cambio se debe contribuir con contenidos propios a RePEc siempre que sea posible y (4) no infringir ninguna declaración de copyright definida por los participantes.
- Su centro de atención se sitúa en los productores y no en los consumidores de información. Es donde mayor diferencia existe con otras iniciativas: todos los servicios RePEc están orientados a proporcionar valores añadidos a los autores o productores de los documentos. Así por ejemplo: análisis de citas, ranking de uso y citas de las publicaciones, páginas personales con documentos publicados y relaciones entre ellos, etc. Es importante señalar que este es el punto en el que están fallando todos los proveedores de servicios de carácter general a la hora de atraer a los investigadores y demostrarles su utilidad. No se está sabiendo demostrar que son útiles y que pueden proporcionarles algún tipo de ayuda en su trabajo profesional. Situación que se agrava cuando se intenta imponer un mandato de archivo sin ofrecer ninguna contraprestación a cambio.
- Su modelo de negocios está basado sobre los mismos principios que mueven la comunidad de software libre: colaboración voluntaria a través de unos procedimientos comunes para aportar metadatos a un *dataset* común, de libre acceso y reutilizable. En un momento en que otros modelos de negocios, de iniciativas como por ejemplo arXiv, han mostrado problemas de sostenibilidad por el alto coste que supone su mantenimiento, es interesante estudiar el modelo de RePEc porque se ha mostrado sólido y sostenible al basarse en una comunidad de usuarios con unos intereses compartidos.

3.- Modelos de negocios desarrollados por algunos repositorios.

Todos los repositorios temáticos, institucionales y recolectores utilizan un modelo de negocios similar, en el que la institución desarrolladora, generalmente una universidad, asume en exclusiva los costes de generación de metadatos, desarrollo de servicios, difusión de contenidos, almacenamiento de documentos, etc. Esto ha sido viable hasta la fecha porque se trata de iniciativas pequeñas, con un reducido número de contenidos y con una trayectoria aun limitada. A medida que los repositorios crezcan esta situación tendrá que revisarse puesto que la institución proveedora puede tener que llegar a hacer frente a unos costes elevadísimos que pongan en cuestión su sostenibilidad a largo plazo.

Aunque no existe mucha literatura al respecto, algunos estudios han contabilizado los costes que deben afrontar las instituciones en el desarrollo de repositorios.

El proyecto LIFE2 hizo un análisis de los costes de repositorios institucionales y recolectores en el Reino Unido (Ayris, 2008). En concreto, analizó el caso del proyecto SHERPA DP¹⁴ financiado por el JISC y RLUK en el período 2005-2007. Este proyecto desarrolló una arquitectura similar a la de RePEc, con diez instituciones ejerciendo de proveedores de datos y una más jugando el papel de proveedor de servicios. En el análisis de costes se tuvieron en cuenta los conceptos identificados en el modelo LIFE v1.1 (Ayris, 2008), si bien, muchos de ellos no fueron aplicables a este caso puesto que se refieren a proveedores de datos. De aquellos que si lo fueron, los que se llevan una mayor parte del presupuesto son: ampliar los metadatos proporcionados por los proveedores debido a que son insuficientes para el funcionamiento del servicio; proceso de recolección; los relacionados con la infraestructura tecnológica tales como asesoramiento sobre requisitos de almacenamiento, sistemas de seguridad, gestión de almacenamiento duplicado o mirror, hardware de almacenamiento y su mantenimiento; soporte a los usuarios y provisión de acceso a los contenidos. Lo más curioso de todo es que este último punto no llegó a desarrollarse, simplemente se puso en funcionamiento un prototipo, con lo que el proyecto no ha dado acceso ni a un sólo documento.

Con una colección mínima de unos 6.500 documentos recolectados LIFE2 estimó que los costes de un proveedor de servicios durante el primer año ascenderían a 143.000 euros o lo que es lo mismo unos 22 euros por

¹⁴ <http://www.sherpadp.org.uk/>

documento. Dado que el escenario descrito se aproxima mucho a RePEc podríamos extrapolar los datos a alguno de sus proveedores de servicios. Así tendríamos que el coste de un portal como EconPapers¹⁵ o IDEAS¹⁶ se situaría en unos 20 millones de euros anuales (942.000 * 22).

Otro objeto de estudio fueron los costes relacionados con la gestión de los proveedores de datos, donde se estimó que el coste de mantener un pequeño archivo de datos se situaría en torno a los 92.000 euros anuales. También LIFE2 examinó los costes de gestionar un repositorio institucional tomando como fuente de su estudio tres casos concretos de universidades del Reino Unido situó el coste entre los 18 y los 37 euros por ítem almacenado.

Ya en el ámbito de los repositorios temáticos, NIH (Nacional Institutes of Health, USA) (NIH, 2008) estima que costaría unos 3.400.000 euros por año mantener los 80.000 artículos que estimaban tenían en 2008, o lo que es lo mismo unos 42 euros por ítem almacenado.

Entre los archivos temáticos el que dispone de una mejor documentación sobre su modelo de negocios es arXiv (Cornell University Library, 2010). Estima que su presupuesto se sitúa en unos 300.000 euros subiendo hasta los 380.000 en 2012. Teniendo en cuenta que durante 2009 se remitieron al archivo unos 64.000 documentos, esto implica unos costes de entorno a los 5 euros por documento, cifra considerablemente menor de las estimadas en los casos anteriores.

En cualquier caso las cifras enumeradas superan las posibilidades de una única institución y, como consecuencia, es necesario buscar nuevos modelos basados en la colaboración.

Este cambio ya se ha producido en dos casos concretos. El primero, se refiere al proveedor de servicios OAster que fue desarrollado inicialmente por la Universidad de Michigan con una subvención de la Andrew W. Mellon Foundation. El año pasado OAster fue cedido a OCLC con objeto de asegurar su viabilidad¹⁷. En segundo lugar, tenemos el caso del repositorio temático arXiv, que fue desarrollado como proyecto en el laboratorio de Los Alamos (USA) y posteriormente transferido a la Universidad de Cornell. El año pasado arXiv cambió su modelo de negocios para buscar financiación entre las instituciones que mayor uso hacen de sus contenidos.

De esta forma establecen tres niveles de participación basándose en el número de descargas desde cada institución a lo largo de un año. Las 100 instituciones con mayor número de documentos descargados pagarán unos 3000 euros al año. Las siguientes 100 unos 2400 y el resto, hasta alcanzar un 20% de todas las descargas institucionales, unos 1700 euros. Aunque se trata de unas cantidades muy reducidas este modelo vuelve a repercutir los costes sobre los usuarios finales, de una forma similar al de las suscripciones.

4.- Modelo de negocios de RePEc

Frente a las cifras señaladas en el apartado anterior, llama la atención que el presupuesto de RePEc sea de 0 euros. Como ya hemos mencionado, el modelo de negocios de RePEc es particular y muy diferente a los analizados en la sección anterior. Se trata de un modelo similar al que ha servido para desarrollar el movimiento de software de código abierto (OSS).

Se considera OSS aquel que permite a cualquier persona, con la formación adecuada, acceder al código fuente del mismo para leerlo, redistribuirlo, modificarlo o extenderlo. Frente al código cerrado que se distribuye en formato binario y al que solamente tienen acceso sus desarrolladores, el abierto permite que cualquiera lo adapte, solucione problemas o incluya nuevas funcionalidades a una velocidad y con una calidad mucho mayor. Cada vez son más los sistemas de software que se distribuyen en esta forma: servidores web como Apache, sistemas operativos como Linux o, en nuestro ámbito, sistemas de gestión de contenidos como DSpace o E-Prints son ejemplos de OSS.

¹⁵ <http://econpapers.repec.org>

¹⁶ <http://ideas.repec.org>

¹⁷ <http://www.oclc.org/news/releases/20096.htm>

Podríamos hacer un paralelismo entre RePEc y el OSS atendiendo a distintos aspectos de análisis como el objeto, el concepto de comunidad, la estructura, las motivaciones y los elementos que los componen.

4.1.- Objeto

El **objeto** en OSS está claramente definido, es el propio sistema de software. En RePEc el objeto es mucho más complejo puesto que se trata no sólo de un conjunto de metadatos sino también de los servicios necesarios para que esos metadatos lleguen de forma efectiva a los investigadores y les sean de utilidad en su trabajo.

4.2.-Comunidad de usuarios

Ambos se basan en el concepto de **comunidad de usuarios**. Frente a la comunicación unidireccional que sigue el software tradicional (desarrolladores que trabajan para usuarios) u otros repositorios (servicios que trabajan para investigadores), en el modelo OSS se explota el concepto de comunidad de intereses donde una serie de elementos interaccionan para alcanzar unos objetivos comunes.

Una comunidad de OSS nace cuando alguien elabora un prototipo funcional de software (Xu, 2009). Los usuarios examinan entonces el prototipo y pueden simplemente adoptarlo o trabajar en común para continuar desarrollándolo. Las personas que comienzan el proyecto normalmente se convierten en los líderes y administradores del mismo, después la comunidad crece a medida que más y más personas son atraídas en el proyecto. Los miembros juegan diferentes papeles: (1) los líderes son los encargados de mantener las versiones; (2) los desarrolladores principales llevan a cabo la mayor parte del desarrollo y mantenimiento; (3) los desarrolladores secundarios informan sobre problemas, sugieren nuevas funcionalidades y ocasionalmente contribuyen con código; (4) los usuarios finales, pasivos, que utilizan el software sin contribuir al código pero que son importantes en asegurar la popularidad del mismo. Otra categorización de roles es la que establece (Shah, 2006) quien distingue entre: los participantes a largo plazo caracterizados porque disfrutan con la programación, son los llamados “hobbyists”; y los participantes a corto plazo, que están motivados por una necesidad inmediata de uso del software.

La comunidad RePEc surgió de una forma similar cuando un grupo de personas creó, en torno a 1997, una masa crítica de metadatos y unos servicios al usuario que los explotaban. Siguiendo la clasificación anterior podríamos considerar a estas personas como los líderes “hobbyist” de la iniciativa: Thomas Krichel, José Manuel Barrueco (CitEc¹⁸), Sune Karlson (EconPapers) y Christian Zimmermann (IDEAS).

A partir de aquí la comunidad comenzó a crecer cuando surgieron otros proveedores de servicios y se fueron uniendo más y más proveedores de datos para formar el grupo de tipo 2 o desarrolladores secundarios. Este grupo, que es el más numeroso, más heterogéneo y en el cuál descansa la principal fortaleza de RePEc tiene dos categorías de miembros: los proveedores de metadatos y los proveedores de servicios.

La provisión de metadatos es una tarea poco gratificante por lo que no se puede esperar que alguien haga este trabajo sin recibir nada a cambio. Además es una tarea tan ingente que escapa a las posibilidades del trabajo tradicional de la biblioteca donde la catalogación de las publicaciones descansa sobre un número limitado de personas. Por ello la idea en RePEc es distribuir la tarea lo más posible y entre aquellas personas o instituciones que tengan incentivos para llevarla a cabo. En la ciencia quienes más incentivos tienen para que un trabajo se difunda son claramente sus productores: autor y editor. Así podemos distinguir dos tipos de proveedores de metadatos: las instituciones, departamentos y editoriales que publican revistas o documentos de trabajo y, por otro lado, los investigadores que contribuyen aportando datos de carácter personal.

Existen más de 1200 colaboradores institucionales que van desde pequeños departamentos con una producción de unos pocos documentos al año, hasta grandes organizaciones como el Banco Mundial o la OCDE con cientos de publicaciones. El grupo de investigadores lo forman más de 25000 personas que se han registrado con RePEc a través del RePEc Author Service (RAS¹⁹). Cuando un autor se registra se le ofrece entre otros servicios: una página personal con sus publicaciones, seguimiento del uso y descargas de las mismas, notificaciones sobre citas recibidas, etc. A cambio debe contribuir manteniendo actualizados sus datos de contacto, haciendo un control de autoridades

¹⁸ <http://citec.repec.org>

¹⁹ <http://authors.repec.org>

sobre las distintas formas de su nombre que permita agrupar todas sus publicaciones bajo una forma única o identificando citas que hayan sido obviadas por el sistema. En el caso de las instituciones la provisión de metadatos es el objeto claro de su participación en RePEc; en los autores, los metadatos se generan de forma indirecta y transparente para ellos a medida que ellos trabajan con el portal. Lo que sería un trabajo tedioso para un bibliotecario se distribuye así entre miles de investigadores que lo hacen de forma gustosa pues están contribuyendo a un aumento de su visibilidad en la disciplina.

La provisión de servicios si que es una tarea mucho más gratificante, sobre todo si esos servicios se usan y son útiles. Proveedores de servicios hay tanto de carácter personal como institucional. Entre los primeros está el sistema de alertas NEP (New Economic Papers²⁰) mantenido por Marco Novarese. La colaboración institucional aquí se centra en dos aspectos, la provisión de servicios íntegros como el caso de la Munich University Library que proporciona el servicio de repositorio temático MPRA (Munich Personal RePEc Archive²¹) destinado a recoger trabajos de autores cuyas instituciones no colaboran con RePEc; y el patrocinio de determinados componentes como alojamiento de máquinas. Así, el Instituto Valenciano de Investigaciones Económicas (IVIE) proporciona alojamiento y conectividad para la máquina que desarrolla el índice de citas CitEc. La State University of New York da soporte a NEP, etc.

En OSS promocionar la implicación de los desarrolladores es muy importante. Los líderes deben atraer y retener personas con una motivación muy fuerte. Igualmente deben demostrar liderazgo favoreciendo las relaciones interpersonales, facilitando las comunicaciones, resolviendo conflictos y mejorando la confianza. Como veremos en la sección siguiente, en el caso de RePEc es sencillo atraer a miembros del tipo 2 porque la implicación necesaria es muy poca en relación a los beneficios de visibilidad que se obtienen, no sucede así con los de tipo 1 donde la implicación es mucho mayor y los beneficios no son tan obvios.

4.3.-Estructura organizativa

En cuanto a la **estructura organizativa**, OSS está basado en las contribuciones de desarrolladores sin el soporte de ninguna organización. A pesar de ello, como señala Hertel (2003), hay algunas características que podrían ser consideradas condiciones estructurales para el éxito de un proyecto. Entre ellas, principios de liderazgo participativo combinados con unas responsabilidades claras; una estructura modular que desprecie la complejidad innecesaria; una política de reconocimiento y crédito que no sólo agradece las contribuciones de programación sino también los trabajos más ingratos como la elaboración de documentación; unas normas claras de funcionamiento de la comunidad que son comunicadas en línea; herramientas de comunicación simples pero fiables que están disponibles de forma universal como e-mail, ftp, etc.

RePEc cumple con todos los criterios enumerados: Sin el soporte de ninguna organización, existen unos roles claros con unas responsabilidades definidas, tanto por los líderes como por cada uno de los participantes. Dispone de una estructura modular donde cualquier persona o institución puede introducir un nuevo servicio de valor añadido o una nueva fuente de datos, igualmente la desaparición de alguno de estos módulos no afecta a la funcionalidad general del sistema. Un principio fundamental es hacer el trabajo sencillo para los miembros por eso se huye de herramientas complejas como protocolos específicos, formatos de metadatos complicados, etc. Para participar en RePEc únicamente es necesario disponer de un servidor web y un editor de texto. El reconocimiento se garantiza a todos los niveles: para los contribuyentes tipo 1 es el prestigio que actualmente tienen sus servicios entre la comunidad académica; para los de tipo 2 tendríamos el aumento del impacto y visibilidad de sus publicaciones. Existen unas normas claras de funcionamiento a partir de dos documentos en los que se basa el trabajo: el Protocolo de Guildford²² y ReDIF²³. Finalmente se dispone también de herramientas de comunicación a través de Internet: listas de correo (repec-run), blog (<http://blog.repec.org>), etc.

4.4.-Motivaciones de los miembros

²⁰ <http://nep.repec.org>

²¹ <http://mpra.repec.org>

²² <http://openlib.org/acmes/root/docu/guilp.html>

²³ http://openlib.org/acmes/root/docu/redif_1.html

Un asunto que ha sido ampliamente tratado en la investigación social del OSS es la **motivación** para contribuir a este tipo de proyectos de forma gratuita y cuáles son las recompensas que se espera obtener a cambio. En el trabajo de Hertel (2003) se identificaron siete factores de motivación. Entre ellos: motivos prácticos relaciones con la mejora de las habilidades propias como desarrollador y usuario del software, motivaciones políticas y sociales para dar soporte al software independiente, motivos más hedonistas tales como el disfrutar programando, etc. Bitzer (2010), por su parte, encontró una elevada correlación entre la actividad profesional en OSS y las contribuciones voluntarias a OSS durante el tiempo libre lo que lleva a concluir que los motivos profesionales juegan un papel importante en las actividades relacionadas con OSS. En la revisión que hace Fang (2009) se incluyen como motivos: aprender y disfrutar, reconocimiento y reputación, oportunidades para avanzar profesionalmente, la ideología del software libre, deseo de identidad social, demostrar sus habilidades superiores de programadores.

En el caso de RePEc todas estas motivaciones podrían ser compartidas por los miembros tipo 1, los que hemos calificado de “hobbyist”, mientras que en el caso de los 2 la motivación principal o casi única es mucho más práctica, se trata de obtener un mayor impacto para sus publicaciones. El impacto es el indicador que mueve la actividad investigadora en todas las disciplinas.

4.5.- Costes

Un último punto a analizar es el de los **costes**. En OSS los costes de desarrollo se distribuyen entre los voluntarios que dedican parte de su tiempo libre a crear código y las empresas/instituciones que pagan a sus desarrolladores para producir un software del que esperan obtener algunos beneficios indirectos.

En el caso de RePEc los costes de creación de metadatos se distribuyen entre las 1200 instituciones participantes que ya hemos visto lo hacen gustosamente porque esperan obtener el premio del impacto y la visibilidad al colocar sus publicaciones en unas fuentes de información de uso generalizado en la disciplina. Además lo hacen bien porque saben que cuantos más y mejores metadatos ofrezcan, mejores servicios de valor añadido obtendrán a cambio. Los costes de los servicios, desarrollados por “hobbyist” son asumidos por los propios líderes como parte de su trabajo para la comunidad.

Existen además la figura del patrocinio, donde determinadas instituciones relacionadas con la economía o las bibliotecas proporcionan servicios como alojamiento y conectividad, compra de equipos, etc. Por ejemplo tenemos el caso ya mencionado de CitEc que está alojado en las instalaciones del IVIE y funciona sobre una máquina donada por el proyecto europeo (EconomicsOnline). Si bien RePEc no forma parte de este proyecto, la práctica totalidad de los contenidos de este portal proceden de RePEc.

Todo ello permite que el sistema se desarrolle con un presupuesto de cero euros.

5.- Análisis de fortalezas y debilidades en el modelo RePEc

Si bien el modelo RePEc se ha mostrado robusto y viable en sus trece años de existencia, debemos analizar si en el futuro podrá mostrarse de la misma forma o si por el contrario no es sostenible a largo plazo. Para ello hacemos un análisis de las fortalezas y debilidades que presenta de cara a plantearnos una serie de modificaciones que deberían tomarse en cuenta en el futuro próximo.

1.5 Fortalezas

- La principal fortaleza de RePEc reside en la comunidad de usuarios sobre la que se asienta, la cual asegura un flujo constante de contenidos que permite un crecimiento a un ritmo de unos 18.000 nuevos ítems cada mes. Esto lo hacen posible los miembros que hemos denominado de tipo 2. Hasta el momento ninguna de las 1.200 instituciones que caen en este grupo ha informado sobre costes excesivos que les lleve a considerar el cese de su colaboración. Pero donde mejor se demuestra la condición de comunidad de RePEc es en la participación por parte de los 25000 autores registrados.

- Otra fortaleza viene del concepto de RePEc como biblioteca abierta con una arquitectura modular y distribuida. Ya hemos dicho que todos los contenidos, tanto los aportados por los proveedores de datos como los generados por los servicios tales como estadísticas de uso o análisis de citas, pueden ser reutilizados por cualquier persona o institución, siempre que no sean con fines comerciales. Esto permite que se creen nuevos servicios bajo demanda o que los ya existentes desaparezcan sin afectar al conjunto de la iniciativa. Esto ha ocurrido varias veces a lo largo de la historia de RePEc. Un ejemplo de nuevo servicio creado recientemente es el Munich Personal RePEc Archive²⁴ que se ha convertido a su vez en un repositorio temático propiamente dicho, con un puesto elevado en el “Ranking web de repositorios del mundo”. Por otro lado tenemos ejemplos de desaparición de servicios: WoPEc (1994-2005) fue el primer proveedor de servicios en RePEc. Este desapareció cuando se desarrollaron otros motores mucho más efectivos como EconPapers o IDEAS.

1.6 Debilidades

- La principal debilidad de RePEc es la excesiva concentración de funciones clave en los miembros de tipo 1, los denominados “hobbyist”. Cada uno de ellos gestiona uno o dos servicios claves para RePEc. Cada servicio funciona de forma autónoma e independiente siendo su gestor el único responsable, pero también la única persona que tiene los conocimientos necesarios para su mantenimiento y desarrollo. Lo que nos lleva a considerar que en caso de incapacidad o ausencia de alguno de ellos, peligraría la continuidad del servicio. Hasta el momento esta situación no ha llegado a producirse pero es un claro problema para el futuro. Si bien podemos suponer que la comunidad respondería de la misma forma que se ha hecho en OSS, no se debería esperar a tener una situación límite para la búsqueda de alternativas. En estos momentos existe una lista de procesos básicos en RePEc que asigna a cada uno al menos dos personas con las competencias necesarias para llevarlos a cabo. Ahora bien, la complejidad de los servicios es tal que la segunda persona tiene nociones, pero le llevaría varios meses tomar la tarea en caso de fallo del desarrollador principal.
- Muy relacionado con lo anterior es la falta de documentación sobre procesos en cada uno de los servicios. Cada servicio funciona de forma personal y no existen documentos que permitan a otros miembros realizar una intervención de urgencia en caso necesario. La documentación es tan imprescindible como las copias de seguridad y al igual que éstas solamente se aprecia cuando se la necesita y no se tiene. La documentación es compleja de elaborar y mucho más de mantener y al no tener una aplicación directa e inmediata se tiende a utilizar el poco tiempo disponible en tareas de desarrollo o atención a los usuarios antes que documentación.
- Incapacidad para atraer nuevos miembros de tipo 1. Una función del grupo de líderes es motivar y atraer al grupo a nuevos desarrolladores que aseguren la renovación del equipo inicial. Esto no se ha conseguido en RePEc. Salvo algunas excepciones, el grupo central ha permanecido inalterado en los últimos diez años. Es cierto que el grupo 2 sigue creciendo de forma constante, a un ritmo de 4-5 nuevas instituciones y varias decenas de autores cada mes, pero faltan candidatos para ayudar a gestionar los servicios existentes, añadir nuevas funcionalidades o desarrollar servicios nuevos. Estos candidatos podrían ser de dos tipos: instituciones que deseen implementar nuevos servicios y estén dispuestas a pagar a su personal para llevarlos a cabo, como el caso reciente de Munich University Library o personas con los conocimientos técnicos necesarios y un deseo de colaborar en proyectos de ámbito internacional. Los candidatos naturales procederían del campo de la Documentación (algunos servicios, por ejemplo CitEc, han nacido como proyectos de investigación en esta área) o de la Informática, dado que el reto es gestionar una biblioteca digital. Desgraciadamente los bibliotecarios no disponen de los conocimientos técnicos necesarios para desarrollar estas tareas y los informáticos no tienen los alicientes para participar en una iniciativa que les parece ajena.
- Falta de recursos para desarrollar nuevos servicios o mejorar los ya existentes. Dado que no se ha conseguido aumentar el número de desarrolladores pero los servicios siguen creciendo, la

consecuencia es que los actuales gestores tienen poco o ningún tiempo para añadir nuevas funcionalidades pues deben trabajar en mantener el sistema, arreglar problemas o tratar con los usuarios. Esto nos lleva a la necesidad de buscar modelos de financiación alternativos que permitan pagar por determinados desarrollos. Vías alternativas pueden ser aumentar los patrocinios, la captación de donaciones o la solicitud de proyectos de investigación.

6.- Conclusiones

En este trabajo hemos descrito el modelo de negocios que sigue uno de los principales repositorios temáticos especializado en Economía, RePEc. Se trata de un modelo original y muy diferente de los que siguen el resto de repositorios temáticos o institucionales que, como RePEc, están trabajando para lograr el acceso abierto a la literatura científica. Está basado en la filosofía del movimiento de software de código abierto y toma como premisa la distribución del trabajo entre una comunidad de personas e instituciones que comparten unos intereses comunes.

Se trata de un modelo que a través de una distribución de costes y la colaboración entre una comunidad de más 1200 instituciones y 25.000 investigadores, ha permitido proporcionar servicios avanzados de valor añadido a un coste cero para los usuarios o receptores de los mismos. A diferencia de otras iniciativas similares que a medida que han ido creciendo han tenido que replantearse la premisa del servicio gratuito a los usuarios, RePEc se ha mostrado sostenible a lo largo de sus trece años de existencia.

Esa fortaleza que da la comunidad de usuarios no está exenta de debilidades. La principal es la concentración de tareas y funciones en los líderes del proyecto junto con una falta de renovación del equipo principal sobre el que descansa la gestión de los principales servicios al usuario.

Es necesario captar la participación de nuevos miembros de tipo 1 que sean capaces de mejorar los servicios ya existentes y desarrollar otros nuevos. En este sentido resulta paradójica la falta de participación de bibliotecarios en RePEc. Es chocante constatar cuando miramos los perfiles profesionales de los miembros de RePEc que, excepto contados casos, la participación de los bibliotecarios se reduce a las tareas tradicionales de provisión de metadatos y no a la gestión de los mismos. Cuando estamos hablando de la mayor biblioteca digital en abierto que existe deberían ser los bibliotecarios quienes estuvieran a la cabeza de su gestión. En unos momentos en los que proliferan los masters universitarios sobre contenidos digitales, RePEc podría ser un campo de pruebas idóneo para la aplicación de verdaderas técnicas de documentación digital. El problema está que los bibliotecarios carecen de (y lo que es peor no consideran necesarios) los conocimientos técnicos necesarios (programación, bases de datos, XML, etc) para llevar a cabo la tarea. Si la profesión no cambia de actitud, desde los profesionales a los formadores, la gestión de contenidos digitales la desarrollarán otros. Y ¿qué será del bibliotecario del siglo XXI si no gestiona contenidos digitales?

Referencias bibliográficas

Ayris, Paul et al. *The LIFE2 final project report*. 2008. <http://eprints.ucl.ac.uk/11758/>

Bitzer, Jürgen; Geishecker, Ingo. "Who contributes voluntarily to OSS? An investigation among German IT employees". En: *Research policy*, 2010, n. 39, pp. 165-172.

Cornell University Library. *arXiv business model white paper*. 2010. <http://arxiv.org/help/support/whitepaper>

Fang, Yulin; Neufeld, Derrick. "Understanding Sustained Participation in Open Source Software Projects". En: *Journal of Management Information System*, 2009, v. 25, n. 4, pp. 9-50.

Henty, Margaret. "Ten major issues in providing a repository service in Australian universities". En: *D-Lib Magazine*, 2007, v. 13, n. 5/6. <http://www.dlib.org/dlib/may07/henty/05henty.html>

Hertel, Guido; Niedner Sven; Herrmann, Stefanie. "Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel". En: *Research policy*, 2003, n. 32, pp. 1159-1177.

NIH. *Analysis of comments and implementation of the NIH public access policy*. 2008.
http://publicaccess.nih.gov/analysis_of_comments_nih_public_access_policy.pdf

Krichel, Thomas. "RePEc, an open library for Economics". En: *The Economics and Usage of Digital Library Collections*. Ann Arbor, Michigan, 2000. <http://eprints.rclis.org/14408/>

Shah, S. K. "Motivation, governance, and the viability of hybrid forms in open source software development". En: *Management Science*, 2006, v. 52, n. 7, pp. 1000-1014.

Xu, Bo; Jones, Donald R.; Shao, Bingjia. "Volunteers' involvement in online community based software development". En: *Information & management*, 2009, n. 46, pp. 151-158.