

# ORGANIZACIÓN DE LOS RECURSOS DIGITALES EN REPEC A TRAVÉS DE LOS METADATOS REDIF

## ORGANIZAÇÃO DE RECURSOS DIGITAIS NO REPEC ATRAVÉS DE METADADOS REDIF

Salvador Enrique Vazquez Moctezuma<sup>a</sup>  
Juan José Calva Gonzalez<sup>b</sup>

### RESUMO

**Introducción:** El repositorio disciplinario RePEc (Research Papers in Economics) provee acceso a un amplio número de preprints, artículos de revistas, libros, capítulos de libros y software sobre la temática de las ciencias económicas y administrativas. Este repositorio agrega registros bibliográficos elaborados por diferentes universidades, institutos, editores y autores que trabajan de forma colaborativa siguiendo normas de organización documental.

**Objetivos:** En este trabajo, principalmente, se identifica y analiza el funcionamiento de RePEc, lo que incluye la organización de los archivos, el cual se caracteriza por el uso del protocolo Guildford y las plantillas de metadatos ReDIF (Research Documentation Information Format) propias para la descripción documental.

**Metodología:** Parte de esta investigación se estudió teóricamente en la literatura; otra parte se llevó a cabo mediante la observación de una serie de características visibles en la página web de RePEc y en los archivos de una revista que colabora en este repositorio.

**Resultados:** Se encontró que el repositorio es un proyecto colaborativo descentralizado, además brinda varios servicios derivados del análisis de metadatos.

**Conclusiones:** Se concluye que las plantillas ReDIF y el protocolo de comunicación Guildford son elementos clave para organizar los registros en RePEc, además existe una similitud con los metadatos de Dublin Core.

**Descriptores:** Repositorio. Preprints. edición electrónica. comunicación científica. metadatos.

---

<sup>a</sup> Mestre em Bibliotecologia y Estudios de la Información pela Universidad Nacional Autónoma de México. E-mail: [quique.vazquez@hotmail.com](mailto:quique.vazquez@hotmail.com)

<sup>b</sup> Doutor em Ciências de la Información, Biblioteconomía y documentación pela Universidad Complutense de Madrid. Professor da Universidad Nacional Autónoma de México. E-mail: [jjcg@unam.mx](mailto:jjcg@unam.mx)

## 1 INTRODUÇÃO

El desarrollo de las tecnologías de la información y comunicación han permitido crear nuevos flujos de comunicación científica a través de los preprints, los artículos de revistas, entre otros documentos distribuidos en la red, a su vez han traído consigo una serie de inconvenientes para su organización como: el crecimiento exponencial, la inestabilidad en Internet además de la diversificación de tipos y formatos de documentos. La desorganización de documentos digitales provoca que se ignore la información que existe y en muchas ocasiones es irrecuperable. De ahí que tenemos la necesidad de organizar los documentos para saber que existe y con esto poder dar a conocer a los usuarios lo que se tiene con el propósito de que se satisfagan las necesidades de información.

La aparición del movimiento de acceso abierto a través de la vía dorada (revistas) y la vía verde (repositorios), se han convertido en canales de comunicación complementarios que permiten una mayor visibilidad e impacto de los resultados de investigación. Si nos centramos en los repositorios encontramos que existen dos tipos: institucionales y disciplinarios. El primero alberga documentos producidos por una institución y sus investigadores, mientras que los disciplinarios condensan información de una temática o área del conocimiento específica, por lo tanto, resguardan más información que los repositorios institucionales. No obstante, el Directory of Open Access Repositories (DOAR) lista actualmente 3.316 repositorios, en los últimos años han surgido 200 nuevos repositorios por año. De estos, un 85% son repositorios institucionales, mientras que los temáticos representan apenas un 9%. (LÓPEZ-BORRULL, 2017, p.242). A pesar del crecimiento del número de repositorios siguen siendo pocos los temáticos que concentran bastante información.

Teniendo en cuenta lo anterior, los repositorios temáticos presentan la dificultad de organizar un mayor número de información, de ahí que se opte por crear las instrucciones precisas para que cada autor o institución efectúe el archivado de los trabajos por iniciativa propia. Por ejemplo, en las ciencias de

la documentación existe el repositorio e-LIS (e-prints in Library and Information Science), y en las ciencias económicas está RePEc (Research Papers in Economics), aunque manejan temáticas distintas tienen la peculiaridad de que cada autor selecciona el tipo de documento que desea archivar (preprints, artículos de revistas, libros, capítulos de libros, etc.), después ingresa metadatos para la organización de los archivos. Por esta razón es indispensable saber cómo funcionan los repositorios para concentrar los registros que contienen descripciones bibliográficas.

En el caso del repositorio de RePEc tiene varias virtudes, como el trabajo colaborativo que ayuda a que cada persona tenga sus registros en el sistema el tiempo que desea sin que se violen derechos de autor, sin embargo se encuentra con un problema en la organización de información, debido a las divisiones que habitualmente hacen los economistas de los temas económicos. Todo ello, en general, impide una organización eficaz lo cual reduce la calidad de las bases de datos y empobreciendo de los resultados en las búsquedas de información (EDO-HERNÁNDEZ, 2010, p.621). Ante esta realidad nos cuestionamos lo siguiente: ¿Quiénes colaboran en el repositorio de RePEc? ¿Cuántos archivos aloja el repositorio en su sistema? ¿Cuál es la estructura para organizar los archivos en el repositorio? Entre otras preguntas que tendrán respuesta si se conoce la estructura de organización del repositorio.

## **2. REPEC**

Research Papers in Economics (RePEc) es un repositorio temático en economía y ciencias afines. Hasta el momento tiene un registro superior de 1.900 archivos pertenecientes a 94 países, asimismo contribuyen alrededor de 2.3 millones de unidades de investigación que incluyen 2.800 revistas y 4.500 documentos de trabajo (working paper), a su vez, se han registrado aproximadamente 50.000 autores (REPEC, 2017). El corazón del repositorio es una base de datos bibliográfica descentralizada de preprints, artículos de revistas, libros, capítulos de libros y componentes de software, el cual es organizado y mantenido por académicos voluntarios. Los fundadores de RePEc se refieren a éste como un servicio bibliográfico o una gran base de datos

(ZIMMERMANN, 2013, p.249), también parece una biblioteca digital descentralizada (BARRUECO; KRICHEL, 2000, p.137), o un sistema académico descentralizado con atributos de una biblioteca abierta (KRICHEL, 2008, p.2). Estas tres categorías que delimitan el repositorio se fundamentan en su desarrollo histórico y su propio modelo organizativo, el cual detallaremos enseguida.

El origen de RePEc inicia con el proyecto NetEc (Networked for Economics), el cual fue fundado en 1993 por una organización internacional que contaba con el esfuerzo de varios voluntarios académicos que deseaban mejorar la comunicación de la comunidad de los economistas a través de los medios electrónicos y de Internet. Esta iniciativa de proyecto propicio el surgimiento de BibEc (Bibliography in Economics) y WoPEc (Working Papers in Economics) los cuales operaban de forma centralizada. El primero compiló 40.000 registros de datos bibliográficos entre 1993-1997, el cual contenía información referencial de preprints de economía impresos por universidades. Mientras que el segundo registraba referencias bibliográficas de preprints que se encontraban en la red alcanzando en un periodo de 4 años 5.000 documentos (KRICHEL, 1997, p.2). Teniendo en cuenta la época en que se desarrollaron estos dos ambiciosos proyectos, podríamos afirmar que el manejo de 45.000 registros, necesitaba de muchos recursos humanos, tecnológicos además económicos por lo que implicaba la organización y alojamiento de los registros de forma centralizada.

Gracias a los 45.000 registros que se obtuvieron a finales del año de 1997, propicio enseguida una reunión entre los miembros del proyecto WoPEc en la ciudad de Guildford en Inglaterra. En dicha reunión se discutió la propuesta del protocolo ReDIF y Guildford así como la fusión de BibEc y WoPEc, los cuales constituyeron el nacimiento y corazón de RePEc (BARRUECO; KRICHEL, 2000c, p.228). Sin embargo, la carencia de presupuesto obligo a la descentralización de actividades en cuestión de organización de registros asimismo solicitaron la ayuda de varios centros de investigación que producían los contenidos de economía, donde pedían que enviaran sus descripciones bibliográficas cuando publicaban nuevos

documentos, por lo tanto, cada centro elaboraba sus propios registros bibliográficos con el propósito de intercambiarlos en Internet entre los interesados en esa temática, además ese modelo de colaboración ofrecía que cada centro de investigación sería dueño exclusivo de sus propios datos. (BARRUECO; KRICHEL, 1999, p.3). Sin duda, RePEc es un proyecto pionero en utilizar la web para difundir los resultados de investigación en las áreas económico-administrativas.

Actualmente existen dos posibilidades de colaboración en RePEc: la primera es proporcionando datos, documentos o archivos y la segunda forma se distingue por el ofrecimiento de servicios a los usuarios finales basado en los archivos que brindaron los primeros. Por su parte, Krichel (2008) advierte que es más práctico dejar que cada institución gestione su propia colección a nivel local (descentralizada) y luego hacerlos accesibles en una interfaz común centralizada mediante el portal de RePEc. Sin embargo, lo anterior es posible gracias a las plantillas que facilitan la descripción bibliográfica del contenido, mismo que sigue el estilo denominado ReDIF, así como el protocolo Guildford que sirve para intercambiar datos. En esencia, las plantillas para la descripción y el protocolo de comunicación representan uno de los logros más importantes, que siguen siendo el marco arquitectónico básico de la corriente de RePEc.

A pesar de esto, el repositorio actualmente ofrece información referencial y documentos en texto completo, aunque los colaboradores pueden optar por ofrecer acceso libre a sus publicaciones o cobrar por el contenido. Cabe destacar que cada autor mantiene todos los derechos sobre sus documentos asimismo tiene la posibilidad de retirar sus trabajos en cualquier momento.

La estructura del repositorio de RePEc (2017) se puede dividir en dos partes: la primera parte, son servicios de alojamiento de metadatos y visualización de la información; mientras que la otra parte son servicios del resultado de análisis de los archivos y metadatos que tiene la base de RePEc (Véase la figura 1).

Con respecto a los servicios de alojamiento de metadatos y visualización de la información nos encontramos que existen cinco tipos: RePEc Author Service, MPRA, IDEAS, EconPapers y RePEc Biblio. Ellos se describen

brevemente a continuación.

- RePEc Author Service, permite a los autores registrarse en RePEc, así cada uno ingresa información que permite relacionar variaciones de nombre que usa al firmar sus artículos y obtener notificaciones de nuevas citas que se encuentran registrados en RePEc.
- MPRA (Munich Personal RePEc Archive) Es una alternativa de archivo central (repositorio) cuando las instituciones de los investigadores no están afiliadas y no participan en RePEc.
- IDEAS (Internet Documents in Economics Access Service), proporciona una interfaz al usuario para navegar y buscar revistas, preprints, libros, capítulos de libros y componentes de software en toda la base de RePEc
- EconPapers, proporciona acceso a RePEc con otra interfaz para buscar información de acuerdo al título de la publicación.
- RePEc Biblio, es una selección manual de bibliografía sobre trabajos destacados en economía.

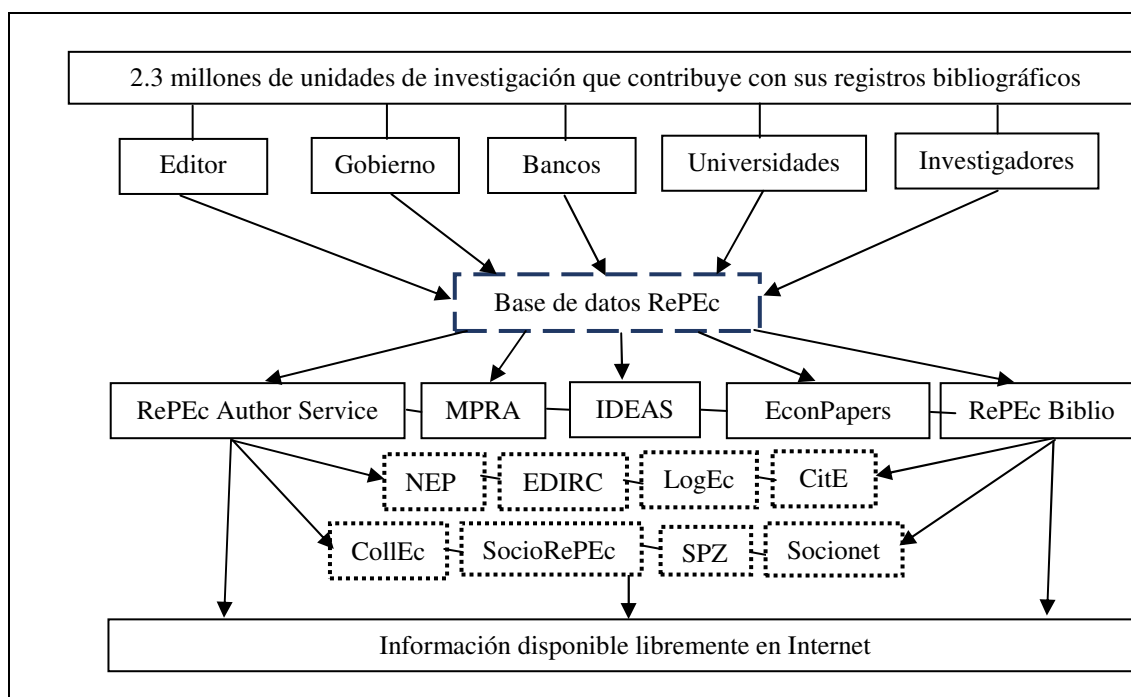
Este repositorio ofrece un conjunto de servicios adicionales a los usuarios, resultado del análisis de los archivos y metadatos que tiene la base de RePEc, entre estos se encuentran: NEP, EDIRC, LogEc, CitEC, CollEc, SocioRePEc.org, SPZ y Socionet los cuales mencionamos a continuación:

- NEP (New Economics Papers), es un servicio de notificación e informes editados sobre los nuevos documentos incorporaciones en RePEc en informes editados. El objetivo de esta herramienta es proporcionar a los suscriptores información actualizada acerca de la literatura de investigación a través del correo electrónico, canales RSS y Twitter.
- EDIRC (Economics Departments, Institutes and Research Centers in the World), es un servicio que indexa a todas las instituciones de economía. Proporciona información detallada sobre la estructura de la institución, sub-entidades como: universidades, departamentos, institutos, y / o centros de investigación, además de una lista de los afiliados. Por otra parte, EDIRC ofrece también una lista llamada Top 25% de las Instituciones y autores de cada país.
- LogEc, es un servicio que proporciona estadísticas de acceso para cada elemento, es decir, ofrece estadísticas generales de la consulta por cada uno de los tipos de documentos además de la descargas de archivos de las interfaces de RePEc.
- CitEC (Citation in Economics), es el servicio que proporciona un índice de citas autónomo de muchos documentos electrónicos distribuidos por RePEc.
- CollEc, ofrece el ranking de centralidad coautoría de autores registrados en el Servicio de Autor de RePEc.
- SocioRePEc.org, es un servicio para registrar documentos RePEc en PDF y establecer relaciones científicas entre las referencias que se encuentren en los documentos.

- SPZ, es un lugar de trabajo en línea para investigadores, tutores y estudiantes dentro del espacio de información RePEc.
- Socionet, es una implementación en idioma ruso para la base de datos en RePEc la cual permite buscar y recuperar información fácilmente.

Son interesantes los servicios que se obtienen a través del análisis de metadatos, por ejemplo, la extracción de citas. RePEc tiene dos fuentes principales para obtener información de citación: En primer lugar, lee todos los documentos disponibles al público dentro de la red, lamentablemente no siempre es posible obtener todas las citas debido a la falta de acceso abierto de los artículos o problemas técnicos. En segundo lugar, los responsables de los archivos (organizaciones o autores) pueden proporcionar información sobre sus citas de sus documentos (SEILER; WOHLRABE, 2012).

**Figura 1. Estructura del repositorio de RePEc.**



Fuente: Elaboración propia.

La variedad de los servicios prestados por RePEc probablemente sean parte de su éxito, debido a que no es un simple repositorio con documentos especializados para la comunidad científica de economía. En el año 2000 Barrueco y Krichel (2000b, p.2) mencionaban que el repositorio carecía de la plantilla de registro bibliográfico de libros y también no tenían estadísticas globales de la base de datos de RePEc, sin embargo, actualmente estas

limitantes han sido superadas gracias al desarrollo del servicio LogEc y las plantillas de libros y capítulos de libros.

Son ya casi dos décadas del surgimiento de RePEc y se puede confirmar que se han mantenido a la vanguardia y actualizándose constantemente. Para ello, podemos observar en la Tabla I, la evolución que ha tenido RePEc con el total de documentos que alberga desde 1998 hasta 2017 y el número de descargas de texto completo por cada tipo, es decir, preprints, artículos de revista, libros, capítulos de libros y aplicaciones de software.

**Tabla I. Tamaño y descargas por tipo de documentos: 1998-2017.**

Año	No. de preprints	Descarga preprints	No. de artículos	Descarga de artículos	No. de libros	Descarga de libros	No. de cap. de libros	Descarga de cap. de libros	No. de software	Descargas de software
1998	54,954	132,139	8,414	3,175	0	0	0	0	261	2,456
1999	61,972	181,014	14,484	14,793	0	0	0	0	474	685
2000	79,650	340,278	34,741	42,694	0	0	0	0	649	6,461
2001	94,562	900,865	52,166	83,621	0	0	0	0	769	27,235
2002	106,257	1,361,162	67,154	317,398	1	56	4	95	856	39,675
2003	120,058	2,254,689	88,461	691,831	409	5,592	338	3,747	1,055	55,596
2004	143,530	2,488,805	152,338	1,148,906	464	9,838	464	6,232	1,513	65,239
2005	169,727	3,020,026	184,993	1,905,807	600	23,164	1,013	9,032	1,333	111,754
2006	199,152	3,372,619	247,840	2,080,334	926	31,836	1,705	49,887	1,456	91,501
2007	224,744	3,706,939	327,430	2,812,398	1,057	40,663	1,865	57,891	1,558	110,475
2008	277,737	4,016,463	399,289	3,558,790	2,338	40,584	2,795	76,221	1,679	123,587
2009	317,488	4,817,245	488,410	4,245,888	5,729	47,239	11,860	129,388	1,871	146,230
2010	382,365	4,743,589	586,492	3,891,937	8,095	54,970	13,448	152,040	2,059	147,191
2011	434,702	3,982,142	682,427	3,925,736	13,317	56,111	14,913	127,794	2,598	143,454
2012	486,675	3,451,482	799,515	3,251,904	13,369	50,035	16,358	106,990	2,879	144,244
2013	554,415	3,436,999	897,249	3,072,343	17,562	54,716	18,772	108,011	3,120	147,386
2014	608,440	3,226,343	1,043,504	2,717,549	19,464	53,086	19,954	98,139	3,365	127,112
2015	705,616	2,927,277	1,203,981	2,516,231	28,425	68,367	36,866	111,019	3,589	142,685
2016	749,600	2,937,377	1,390,841	2,341,244	31,092	77,733	43,376	117,160	3,803	137,079
2017	794,125	2,798,469	1,591,341	2,455,926	34,637	84,585	46,918	127,942	4,028	129,451

**Fuente:** Elaboración propia con base a las estadísticas de RePEc; Disponible en: <http://logec.repec.org/details.htm>

Es interesante ver, con respecto a los diversos tipos de documentos como inicio RePEc en 1998 concentrando más preprints, seguido de artículos de revistas y aplicaciones de software, pero fue hasta el año 2002 que agregaron libros y capítulos de libros. Otro momento destacable es el año 2004, pues RePEc registró 152,338 artículos de revista y 143,530 preprints lo que denota un incremento en el número de artículos alojados en el repositorio y



posible consecuencia a los efectos del movimiento del acceso abierto.

Actualmente podemos distinguir que las aplicaciones de software son el menor número de registros, pero entre los libros y capítulos de libros señalamos que hay más capítulos de libros desde el 2005. Así mismo, la cantidad de artículos ha crecido sorprendentemente y rebasa a los preprints. De forma general identificamos que desde 1998 hasta 2017, los preprints ha sido el material más descargado, seguido de los artículos de revista, continuado las aplicaciones de software, después los capítulos de libros y por último los libros. Sin duda, las estadísticas de RePEc ofrecen información de su historia por lo cual intuimos que en el futuro siga siendo uno de los principales repositorios especializados.

### **3. ARQUITECTURA DE LOS ARCHIVOS EN REPEC**

La estructura básica de RePEc se puede resumir en tres partes: Archivos, core site y servicios (BARRUECO; KRICHEL, 1999, p.4). Aunque gran parte de su funcionamiento es gracias al uso de las plantillas de descripción documental con el formato de metadatos ReDIF (Research Documentation Information Format) para la organización del contenido y el protocolo Guildford para el intercambio de datos. Todos estos elementos los describimos en seguida.

1. Archivos: Los archivos son proporcionados por los departamentos o autores. Para ello cada uno de forma independiente realiza la descripción de sus documentos (libros, capítulos de libros, preprints, artículos y programas de software) a través de unas plantillas que están ya definidas por RePEc, es decir, los archivos que se hacen, contienen metadatos los cuales se comparten en internet.
2. Core site: A pesar de que RePEc trabaja de forma descentralizada los registros es necesario recopilar toda la información de los colaboradores, para ello tiene un sistema central el cual se puede ver en la página web <http://repec.org/>, sí realizamos una consulta a su servidor podemos confirmar que la dirección IP es: 136.167.18.142 misma que pertenece al Boston College, ubicada en Chestnut Hill, Massachusetts. Gracias a esta verificación de IP podemos decir que el Core site de RePEc se encuentra en esta universidad.

Por su parte, los fundadores de RePEc (BARRUECO; KRICHEL, 1999, p.5) mencionan que para mantener el sistema son indispensables otras funciones como:

- El mantenimiento y actualización de la documentación, el protocolo, las páginas del servidor web, guías actualizadas con información para los colaboradores, las especificaciones de las plantillas *ReDIF*, las cuales explicamos más adelante.
  - La inclusión y retiro de archivos del sistema.
  - La asignación de códigos de identificación a cada nuevo archivo. Cualquier departamento que desea colaborar con RePEc se le asigna una clave compuesta de tres letras.
  - El software es mantenido y actualizado por el archivo central, de forma que cualquier departamento, aunque carezca de personal informático, pueda crear un archivo o servicio. El lenguaje de programación utilizado es *Perl*, y los programas han sido diseñados inicialmente para funcionar en máquinas *Unix*.
3. Servicios: Los archivos que proporcionan los departamentos y autores a través del uso de los metadatos ReDIF de forma separada no tienen ningún valor, aunque en su conjunto permiten crear servicios y así darles un valor agregado debido al análisis de metadatos que realizan las mismas entidades que colaboran con sus archivos que únicamente se dedican a crear algún servicio. Por su parte, Subirats y Barrueco (2002, p.144), mencionan que aquel valor agregado puede tomar distintas formas: la más sencilla sería convertir las descripciones bibliográficas del formato ReDIF a otro legible, por ejemplo, HTML o XML. También se puede consistir un índice de los datos que permita realizar búsquedas y recuperación de la información. Otra opción sería realizar una selección, comentario y presentación de determinados documentos existentes en el sistema.

La arquitectura de RePEc permite trabajar de forma distribuida los archivos que después se concentra en el Core site y son disponibles en la red, además ofrece la posibilidad de crear servicios con los archivos almacenados, es decir, darles un valor agregado a los registros bibliográficos. Sin embargo, para poder hacer posible este funcionamiento es necesario el uso del protocolo Guildford, el cual propicia el intercambio de información en la red, asimismo la utilización de los metadatos ReDIF (Research Documentation Information Format), los cuales favorecen en la descripción de los documentos como: preprints, artículos de revistas, libros, capítulos de libros y aplicaciones de

software.

### 3.1 EL PROTOCOLO GUILDFORD

En los sistemas informáticos son básicos los protocolos de comunicación, mismos que tienen la función de determinar las reglas para efectuar la transmisión de información entre un sistema. Por su parte RePEc ideó el protocolo Guildford para regular los procesos comunicativos de publicación e intercambio de metadatos en internet; sin el seguimiento de dicho protocolo sería imposible la comunicación.

De acuerdo con Krichel, (1999, p.1), la idea detrás del protocolo Guildford se remonta a una declaración de William L. Goffe, administrador de NETEC en 1995, que dice:

“Lo que yo sugiero es esto: un sistema distribuido con cualquier número de sitios, cada uno reflejando unos a otros. Tendría amplias funciones bibliográficas (referencias cruzadas, etc.), y mis favoritos, marcas de tiempo digitales para cuando los documentos fueron presentados. Para los archivos fuera de ella, los documentos podrían enumerarse, pero sin referencias cruzadas. Pero, esos archivos podrían "unirse" el sistema (decir que fue escrito en Perl así que podría ejecutarse en NT y Unix). Entonces usted tiene lo mejor de ambos mundos: distribuida, cualquiera podía unirse, extensa referencias cruzadas, todo funciona. Este sistema podría crecer fácilmente con el uso de la profesión de la red. Dicho sistema sería de gran beneficio de la profesión” (KRICHEL, 1999, p.1).

El protocolo Guildford, cuyo nombre hace referencia a una ciudad del sur de Inglaterra donde fue creado. En esencia este protocolo ofrece un conjunto de reglas para la publicación e intercambio de documentos y metadatos a través de Internet. Además, se caracteriza porque puede ser implementado por cualquier institución o persona que desee intercambiar documentos en la Red (SUBIRATS; BARRUECO, 2002, p.145). Para ello, establece dos niveles en la participación de los departamentos: 1) el de archivo, de carácter pasivo pues simplemente proporciona información y 2) el de servicio de carácter activo, ya que extrae de los archivos y construye un servicio de utilidad para los usuarios finales (BARRUECO; KRICHEL, 1999, p.5). Por lo tanto, ignorar su uso haría imposible la comunicación entre las distintas computadoras y servidores, al menos en el sistema de RePEc.

Para regular la comunicación entre un archivo y los servicios, naturalmente se necesita un archivo que contiene un conjunto de directorios y subdirectorios accesibles a través de un servidor ftp o http, lógicamente la parte más importante del protocolo es aquella donde se describe tal estructura. De acuerdo con Barrueco y Krichel advierten que todo archivo debe ser identificado en el sistema por un código único (`archivo_id`), que es asignado por el core site y compuesto por tres letras. Este aspecto constituye el elemento raíz de la estructura. Los archivos, a su vez, determinarán las distintas colecciones o series de documentos que publiquen, esta vez de seis letras (`series_id`). Cada serie dispondrá de un subdirectorio donde se almacenarán las descripciones bibliográficas en texto plano. Sin embargo también son necesarios dos ficheros que describen al departamento que colabora con sus archivos. Así dentro del archivo todos los ficheros tienen extensión `.rdf` los cuales contendrán los metadatos en formato ReDIF (BARRUECO; KRICHEL, 1999; 2000d). Aquí es indispensable señalar que la extensión `.rdf` no hace referencia al sistema de metadatos RDF, sino más bien es la extensión del fichero necesaria para ejecutarlo o interpretarlo por los sistemas utilizados en RePEc.

### **3.2. LOS METADATOS REDIF**

Una alternativa para organizar y localizar documentos digitales es a través de los metadatos; este término, fue acuñado por Jack Myers en la década de los sesenta para describir conjuntos de datos; el significado que se le dio fue dato sobre un dato, ya que proporciona la información mínima necesaria para identificar un recurso, además de incluir información descriptiva sobre el contexto, calidad y condición o características del dato (SENSO; ROSA, 2003, p.95), por lo tanto, los metadatos son también, igual que la descripción bibliográfica, una forma de organizar la información para su recuperación (PÉREZ, 2006, p.2). Fundamentalmente los metadatos funcionan como elemento de enlace en la búsqueda en tanto que describen el contenido y localización de la información, una función muy parecida a la de los catálogos con la distinción del formato del documento y del procedimiento automatizado.

Existen diversos sistemas de metadatos como: Dublin Core, Text Encoding Initiative, Resource Description Framework, entre otros. Los metadatos ReDIF del acrónimo de Research Documentation Information, se inspiran en el IAFA (Internet Anonymous Ftp Archive), el cual toma la idea de crear plantillas con un conjunto de campos permitidos, algunos de ellos obligatorios, y otros se pueden repetir, así como el uso de clusters (KARLSSON; KRICHEL, 1999, p.205). Sin duda el éxito de ReDIF es gracias a la sintaxis de una plantilla, pues es más legible y fácil de entender por el ser humano.

El concepto de clusters, son los datos referidos a un mismo objeto que deben aparecer juntos en la descripción y pueden ser definidos de la siguiente forma: existe una serie de datos que son comunes en varios casos. Por ejemplo, cada persona mencionada en un registro puede tener un nombre, un lugar de trabajo, una dirección de correo electrónico, etc. De igual forma, cada organización tendrá también información concreta para cada una de esas categorías. A ese conjunto de elementos, los que son comunes a ambos, es a lo que se llama cluster. Así, sería un grupo de campos enlazados lógicamente en un registro y que se refieren a un objeto determinado y que deben aparecer seguidos dentro del registro (BARRUECO; KRICHEL, 1999, p.4). En esencia los cluster son una relación de entidades y puntos de acceso clave.

Los metadatos ReDIF, son un formato de datos en texto plano, que permite un único fichero que pueden almacenar varios registros del mismo tipo, y cada uno debe ocupar un párrafo de texto, es decir, no están permitidas líneas en blanco en su interior (KRICHEL, 2007). El inicio de la plantilla comienza con un campo obligatorio Template-type cuyo contenido es alguno de los siete tipos mostrados en la tabla II. En el resto de los campos no importa el orden en que aparezcan.

**Tabla II. Plantillas ReDIF y campos de descripción**

<i>Nombre del Template-type</i>	<i>Descripción de documento</i>	<i>Campos requeridos</i>	<i>Campos adicionales</i>
<i>ReDIF-Archive 1.0</i>	<i>Es puramente de naturaleza administrativa y sirve para informar RePEc de la ubicación de los</i>	<i>Handle Name Maintainer-Email Description URL</i>	

	<i>archivos de metadatos y proporciona un contacto (técnico) si se desarrolla un problema.</i>		
<i>ReDIF-Series 1.0</i>	<i>La plantilla es utilizada por los servicios y mostrar información acerca de su serie.</i>	<i>Name Provider-Name Provider-Homepage Maintainer-Name Maintainer-Email Type Handle</i>	<i>Provider-Institution Order-Email Price Restriction Maintainer-Phone Editor-Name Notification ISSN Predecessor</i>
<i>ReDIF-Paper 1.0</i>	<i>Working papers o preprints</i>	<i>*Author-Name *Author-Name-First *Author-Name-Last *Author-Email *Author-Workplace-Name Title Abstract Length Creation-Date Revision-Date Publication-Status *File-URL *File-Format *File-Function Number Classification-JEL Keywords Handle</i>	<i>File-Restriction Note Article-Handle</i>
<i>ReDIF-Article 1.0</i>	<i>Artículos publicados en revistas</i>	<i>*Author-Name *Author-Email *Author-Workplace-Name Title Abstract Classification-JEL Keywords Journal Pages Volume Issue Year Month *File-URL *File-Format Handle</i>	<i>File-Restriction Number Paper-Handle</i>
<i>ReDIF-Book 1.0</i>	<i>Libros</i>	<i>*Editor-Name *Editor-Email *Editor-Workplace-Name Title Provider-Name Abstract Classification-JEL Keywords HasChapter Volume Edition</i>	<i>File-Restriction Note Paper-Handle Article-Handle Software-Handle ISBN</i>

		<i>Year</i> <i>Month: February</i> <i>*File-URL</i> <i>*File-Format</i> <i>Order-URL</i> <i>Handle</i>	
ReDIF-Chapter 1.0	<i>Capítulos de libros</i>	<i>*Author-Name</i> <i>*Author-Email</i> <i>*Author-Workplace-Name</i> <i>*Editor-Name</i> <i>*Editor-Email</i> <i>*Editor-Workplace-Name</i> <i>Title</i> <i>Book-Title</i> <i>Abstract</i> <i>Classification-JEL</i> <i>Keywords</i> <i>In-Book</i> <i>Pages</i> <i>Volume</i> <i>Chapter</i> <i>Edition</i> <i>Year</i> <i>Month</i> <i>*File-URL</i> <i>*File-Format</i> <i>Handle</i>	<i>File-Restriction</i> <i>Note</i> <i>Paper-Handle</i> <i>Article-Handle</i> <i>Book-Handle</i> <i>Software-Handle</i>
<i>ReDIF-Software 1.0</i>	<i>Aplicaciones de software</i>	<i>Title</i> <i>*Author-Name</i> <i>*Author-WorkPlace-Name</i> <i>*Author-Email</i> <i>Abstract</i> <i>Series</i> <i>Number</i> <i>Creation-Date</i> <i>Length</i> <i>Classification-JEL</i> <i>*File-URL</i> <i>*File-Format</i> <i>*File-Function</i> <i>Handle</i>	<i>File-Restriction</i> <i>Paper-Handle</i> <i>Article-Handle</i> <i>Book-Handle</i> <i>Chapter –Handle</i> <i>Note</i> <i>Requires</i> <i>Programming-language</i>
<i>* Clusters</i>			

**Fuente:** Elaboración propia con base en las plantillas ReDIF de RePEc, disponible en: <https://ideas.repec.org/trdfintro.html>

Como se observa en la Tabla II, las plantillas ReDIF son un sistema estructurado por campos o etiquetas de metadatos, aunque se les debe agregar dos puntos y un espacio para describir el contenido del documento o valor. En su arquitectura se tiene bien establecidos los campos obligatorios como: *Template-Type*, *Author-Name*, *Editor-Name*, *Title*, *Provider-Name*, y *Handle*, contenidos en las plantillas. El handle es un elemento clave en todas las plantillas ReDIF, y es una cadena de caracteres que representan

información del archivo, serie y objeto, permitiendo así una relación de entidades. Es de notarse que los campos de metadatos permiten una flexibilidad y descripción necesaria para cada tipo de documento, para ello las plantillas ReDIF son de gran ayuda cuando se carece de personal especializado en ciencias de la información o tecnologías de la información lo que facilita la descripción documental de forma organizada, aunque se puede completar los registros con campos adicionales, es decir entre más datos se ingresen mayores serán las alternativas para recuperar la información.

Por otro lado, en el entorno de la documentación se podría debatir sobre la creación del formato de metadatos ReDIF y cuestionar porque el repositorio de RePEc no utiliza alguno de los sistemas ya consolidados como: Marc o Dublin Core. De acuerdo con Barrueco y Krichel (1999; 2005), advierten que las principales razones de utilizar ReDIF es por la sencillez que ofrece para la descripción bibliográfica, lo que permite elaborarlas por personal no especializado, considerando que Marc a veces es difícil de entender por los propios bibliotecarios. Otro punto es que los metadatos de las plantillas ReDIF permite la actualización de forma rápida. También otro aspecto fundamental es causado por razones históricas, el número de registros de WoPEc que usaban el formato IAFA imposibilitaban la sustitución por otro sistema emergente, pues los metadatos de Dublin Core estaban apenas en desarrollo y pocos sistemas de búsqueda interactuaban con esos metadatos. Aunque los metadatos ReDIF están diseñados para la comunidad de los economistas, es posible modificarlos para su implementación en otros repositorios temáticos con poca infraestructura gracias a la flexibilidad que tiene (KRICHEL; ZIMMERMANN, 2009, p.143). Además han participado bibliotecarios en el diseño de los metadatos ReDIF por lo que considera aspectos mínimos para la descripción bibliográfica óptima de los documentos y permite extraer los metadatos para generar algún servicio.

Sin duda, para que funcione el repositorio de RePEc es necesario seguir las reglas que establece el protocolo Guildford, es decir, los elementos (ficheros, directorios, etc.) necesarios para poder comunicar los datos a través de la red con el uso de un servidor ftp o servidor web. Asimismo los elementos



contienen mínimo tres plantillas ReDIF como el de serie, archivo y la del tipo de documento a describir (preprints, artículo de revista, libro, capítulo de libro y software), estas plantillas son esenciales porque permiten el trabajo descentralizado, además sirven en la búsqueda y recuperación de información en el sistema de RePEc por parte de los usuarios finales. Gracias a las bondades del protocolo y las plantillas ReDIF hacen posible el funcionamiento entre los archivos que proporcionan los colaboradores o instituciones de economía que son compartidos al core site a través de la url del servidor y a su vez el core site comparte los metadatos recabados para prestar servicios. La organización de los documentos que presente el repositorio RePEc se vislumbra por las plantillas ReDIF y el protocolo.

#### **4. COLABORACIÓN DE LA REVISTA CONTADURÍA Y ADMINISTRACIÓN EN REPEC**

El siguiente apartado es una visión sobre como la revista *Contaduría y Administración* publicada por la Universidad Nacional Autónoma de México participa en la colaboración de sus archivos en RePEc, además se señalan las principales características del uso de metadatos. Actualmente RePEc tiene registradas, para el caso de México, 75 organizaciones que incluyen bancos, asociaciones, universidades, centros de investigación entre otros departamentos que describen sus documentos.

*Contaduría y Administración* es una revista trimestral arbitrada por pares, cuyo objetivo es contribuir al avance del conocimiento científico y técnico en los campos de la contaduría, la administración y la informática, mediante la publicación de artículos de investigación teórica o aplicada, incluyendo aquellos que analizan la problemática que enfrentan las tareas de investigación en dichas disciplinas. Desde el año 2014 la revista se encuentra colaborando con RePEc lo que le ha brindado una mayor visibilidad, para ello cuando se inició el proceso para trabajar con RePEc, se siguieron varias instrucciones para el registro, llevándonos a utilizar un instructivo de cómo ir paso a paso en la colaboración de archivos, (<https://ideas.repec.org/stepbystep.html>). Desde la perspectiva de la revista como colaboradora de archivos podemos decir que

RePEc necesita un conjunto de directorios alojados en un servidor de tipo FTP, con un directorio raíz denominado “nax” el cual fue definido por RePEc, dentro de este directorio encontramos en su interior tres elementos (véase figura 2) y a continuación describimos su forma:

- Hay un documento llamado “naxarch.rdf”. Este fichero es indispensable en todo archivo, así mismo utiliza la plantilla denominada archive templates, la cual está disponible en la página de RePEc en la sección Prepare metadata for your RePEc archive. Como se observa abajo, aquí se localiza información de la institución colaboradora, nombre y email del contacto de mantenimiento de los archivos, una breve descripción de la unidad colaboradora, la URL donde se localizan los ficheros, también se detalla la relación del identificador de tres caracteres “nax” con el colaborador.

```
Template-type: ReDIF-Archive 1.0
Handle: RePEc:nax
Name: FCA, Universidad Nacional Autónoma de México
Maintainer-Email: revista_cya@fca.unam.mx
Description: Contaduría y Administración, es una revista arbitrada...
URL: ftp://132.248.18.227/RePEc/nax/
```

- Mientras que el archivo naxseri.rdf, de la misma forma que el anterior fichero es obligatorio y utiliza la plantilla conocida “The series template”. Aquí se representa la información del equipo de trabajo que crea los archivos además se encuentra el nombre de las series, su descripción, los códigos de identificación, el tipo de documento que se describirá, etc. Esta estructura debe seguir reglas establecidas en la plantilla para permitir a los robots de RePEc el acceso a los datos y en caso de error se es notificado.

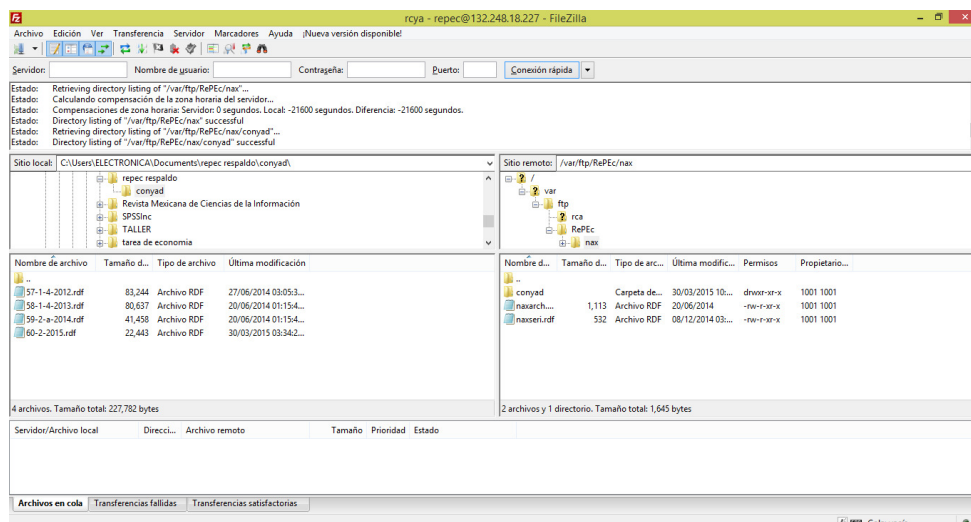
```
Template-type: ReDIF-Series 1.0
Name: Contaduría y Administración
Provider-Name: Accounting and Management
Provider-Homepage: http://www.cya.unam.mx/index.php/cya/
Provider-Institution: RePEc:edi:fcunamx
Editor-Name: Francisco López-Herrera (Director Editorial)
Editor-Email: revista_cya@fca.unam.mx
Maintainer-Name: Salvador Vazquez
Maintainer-Email: revista_cya@fca.unam.mx
ISSN: 0186-1042
Type: ReDIF-Article
Handle: RePEc:nax:conyad
```

- Por otra parte, se tiene un subdirectorío denominado “conyad” (este nombre lo define el colaborador, pero sólo se permite el uso de 6 caracteres), el cual está integrado por archivos de texto plano o ASCII con extensión .rdf, que contiene el registro bibliográfico o metadatos de cada uno de los artículos, como vemos abajo en el registro. Aquí es interesante ver que cada registro de artículo debe iniciar con “Template-Type: ReDIF-Article 1.0” para separarlo de un artículo de otro, además estas plantillas funcionan basadas en la estructura campo: valor, de ahí que después de los dos puntos debe existir un espacio y enseguida la información, así mismo cada campo debe estar en un renglón.

```
Template-Type: ReDIF-Article 1.0
Author-Name: Raúl de Jesús Gutiérrez
Author-Email: rjg2005mx@yahoo.com.mx
Author-Workplace-Name: Universidad Autónoma del Estado de México,
México
Author-Name: Edgar Ortiz
Author-Workplace-Name: Universidad Nacional Autónoma de México,
México
Author-Name: Oswaldo García Salgado
Author-Workplace-Name: Universidad Autónoma del Estado de México,
México
Title: Los efectos de largo plazo de la asimetría y persistencia en
la...
Abstract: Este trabajo propone una extensión al modelo CGARCH ...
Classification-JEL: C22,C32,C51,C52
Keywords: Volatilidad asimétrica; Mercados accionarios emergentes;
Pages: 1063-1080
Volume: 62
Issue: 4
Year: 2017
Month: Octubre-Diciembre
File-URL: http://www.cya.unam.mx/index.php/cya/article/view/974/906
File-Format: Application/pdf
Handle: RePEc:nax:conyad:v:62:y:2017:i:4:p:1063-1080 File-Format:
Application/pdf
Handle: RePEc:nax:conyad:v:60:y:2015:i:2:p:291-319
```

Técnicamente se puede ver complejo la creación de directorios y uso de plantillas para la creación de los ficheros, sin embargo la complicación es solo al inicio, porque se debe definir que campos se van a utilizar en los archivos que estarán en el subdirectorío de seis caracteres y como se presentará la información u orden. Con respecto a los ficheros naxarch.rdf y naxseri.rdf, es información que es fácil completar por los colaboradores además que deben coincidir los nombres de los ficheros y handle para que exista una relación y comunicación entre la unidad colaboradora y los archivos que ofrece.

Figura 2. Estructura de los archivos que elabora una revista en RePEc.



En la figura 2, observamos los archivos creados por la revista Contaduría y Administración para compartir con RePEc. Con una interfaz de Filezilla, cliente FTP multiplataforma que facilita la administración de documentos en un servidor, de lado derecho encontramos la organización de los directorios y los archivos que se compartirán a través del servidor FTP con el sistema RePEc. Así tenemos el directorio nax, en él se encuentra el subdirectorio conyad además los ficheros naxarch.rdf y naxseri.rdf visible en: (ftp://132.248.18.227/RePEc/nax/). Mientras de lado izquierdo están los documentos que deseamos copiar para agregarlos en la comunicación en línea, aunque estos archivos son propios de cada computadora y pueden variar, por ello cada fichero representa un volumen publicado por año el cual tiene cuatro números de la revista Contaduría y Administración. Es indispensable que cada fichero que pongamos en el servidor FTP tenga todos los permisos para facilitar la extracción de metadatos por el sistema de RePEc.

Con respecto a los metadatos inmersos en las plantillas ReDIF hemos identificado que personal carente de habilidades sobre organización de la información, se le puede hacer fácil el llenado de una plantilla e ingresar datos sin percibirse que está ingresando metadatos. Asimismo, el sistema de RePEc notifica cuando se tiene un campo con información errónea, señalando la línea donde se debe corregir. También existe una aplicación web que ayuda a validar el registro disponible en: (<https://econpapers.repec.org/scripts/redifcheck.pl>).

De acuerdo a la literatura que revisamos sobre los metadatos ReDIF, mencionaba que no se utilizaba Dublín Core, pero hemos detectado que actualmente RePEc utiliza las principales etiquetas meta DC, es decir Dublín Core. Pudimos encontrar esta característica gracias a que ingresamos a varios registros de artículos dispuestos en el portal de RePEc como un usuario final y en los registros colocamos la opción de ver código fuente de la página, de ahí que observamos metadatos en DC (véase figura 3).

Figura 3. Etiquetas meta Dublín Core en RePEc

```
13 <link rel="stylesheet" type="text/css" href="/EconPapers.css">
14
15 <title>EconPapers: Análisis del riesgo de mercado de los fondos de pensión en México Un enfoque con modelos autorregresivos</title>
16 <!-- BEGIN META -->
17 <LINK REL="SCHEMA.DC" HREF="http://purl.org/dc/elements/1.1/">
18 <META NAME="dc.Title" CONTENT="Análisis del riesgo de mercado de los fondos de pensión en México Un enfoque con modelos autorregresivos">
19 <META NAME="dc.Creator" CONTENT="R., Martínez Preece Marissa">
20 <META NAME="dc.Creator" CONTENT="Francisco, Venegas Martínez">
21 <meta name="dc.Subject" content="Fondos de pensiones, SIEFORES, esquema de jubilación, riesgo de mercado, premio al riesgo">
22 <meta name="dc.Publisher" content="Accounting and Management">
23 <meta name="dc.Type" content="text">
24 <meta name="dc.Type" scheme="DCT2" content="article">
25 <meta name="dc.Date" content="2014">
26 <meta name="redif-type" content="article">
27 <meta name="handle" content="RePEc:nax:conyad:v:59:y:2014:i:3:p:165-195">
28 <meta name="JEL-Codes" content="C22 C58 G23 G32">
29 <meta name="EPDATE" content="2014-06-27">
30 <meta name="description" content="By Martínez Preece Marissa R. and Venegas Martínez Francisco; Abstract: The aim of this paper is to analyze the market risk of two types of investment funds, Basic SIEFORE 1 (SB1) and Basic">
31 <meta name="series" content="Contaduría y Administración">
32 <meta name="citation_authors" content="R., Martínez Preece Marissa; Francisco, Venegas Martínez">
33 <meta name="citation_author" content="R., Martínez Preece Marissa">
34 <meta name="citation_author" content="Francisco, Venegas Martínez">
35 <meta name="citation_title" content="Análisis del riesgo de mercado de los fondos de pensión en México Un enfoque con modelos autorregresivos">
36 <meta name="citation_date" content="2014">
37 <meta name="citation_year" content="2014">
38 <meta name="citation_publication_date" content="2014">
39 <meta name="citation_journal_title" content="Contaduría y Administración">
40 <meta name="citation_publisher" content="Accounting and Management">
41 <meta name="citation_volume" content="59">
42 <meta name="citation_issue" content="3">
43 <meta name="citation_firstpage" content="165">
44 <meta name="citation_lastpage" content="195">
45 <meta name="citation_issn" content="0186-1042">
46 <meta name="citation_keywords" content="Fondos de pensiones; SIEFORES; esquema de jubilación; riesgo de mercado; premio al riesgo">
47 <meta name="citation_abstract" content="The aim of this paper is to analyze the market risk of investment funds, Basic SIEFORE 1 (SB1) and Basic SIEFORE 2 (SB2). To do this, we propose a performance index that will be used in ARIMA-GARCH models and some of its extensions, with the purpose of examining the dynamic behavior of the returns and their volatility in such investment funds. Moreover, the identification of risk types of funds is included. One of the relative strength
```

La identificación del uso de metadatos Dublín Core en RePEc nos llevó hacer una comparación de los 15 campos de DC (HILMANN, 2011), mostrados en la Tabla III y determinar las etiquetas que se usan para organizar la información. En ese comparativo pudimos notar que se utiliza poco Dublín Core en la arquitectura de las plantillas ReDIF, aunque las que se llegan a usar son las más indispensables. De ahí que los registros ReDIF usen los siguientes campos de Dublín Core: Título, creador, materias, descripción, editor, fecha, tipo, formato, identificador y fuente. En la verificación que realizamos nos fue imposible encontrar determinadas etiquetas DC en RePEc, por ejemplo, las etiquetas de colaborador, idioma, relación y derechos. Intuimos que a pesar de que no se establecen esos campos, sí están considerados en las plantillas en ReDIF de RePEc.

**Tabla III. Relación entre los metadatos Dublin Core y ReDIF utilizados en la indexación de documentos mostrados en RePEc.**

1 Dublin Core	2 Descripción de campos en plantillas ReDIF
1. Título	3 Título del preprint, artículo de revista, libro, capítulo de libro y software.
2. Creador	4 Nombre del autor, afiliación y correo electrónico.
3. Materias	5 Palabras clave.
4. Descripción	6 Abstract del documento.
5. Editor	7 Editor, patrocinador o institución de investigación.
6. Colaborador	8 ***
7. Fecha	9 Fecha de publicación.
8. Tipo	10 ReDIF-Paper (documentos de trabajo e investigación (working paper), pre-prints, post-prints), ReDIF-Article (Contenidos de revistas), ReDIF-Software (Codigos de programación y software), ReDIF-Book (Libros), ReDIF-Chapter (Capítulos de libros).
9. Formato	11 File, PDF, html, txt y lenguaje de programación.
10. Indentificador	12 Handle
11. Fuente	13 Título de revista o libro, volumen y número
12. Idioma	14 ***
13. Relación	15 ***
14. Cobertura	16 ***
15. Derechos	17 ***

\*\*\* Elementos que no utiliza RePEc

Por ejemplo, en la tabla anterior notamos que el campo de colaborador sí lo contrastamos con las plantillas de la tabla II, esta sustituida en el campo autor o editor. En cuanto al idioma, RePEc no lo considera en ningún campo. Por su parte, el campo de relación lo encontramos en “Paper-Handle, Article-Handle, Book-Handle, Chapter-Handle y Software-Handle”, los cuales relaciona documentos dentro del repositorio. Con respecto a los derechos de autor, no se maneja algún campo, debido a que la filosofía del repositorio de RePEc es trabajar de manera colaborativa pero descentralizada, entonces cada colaborador decide cuando retirar sus archivos, a pesar de esto es el propio colaborador quien decidirá si cobra por su contenido o lo deja en acceso libre.

Por último, el repositorio de RePEc ofrece a las editoriales mayor

visibilidad, aunque es necesario seguir sus lineamientos en la creación de archivos para compartirlos y ellos los traten para su visualización en Internet. Por otra parte, los metadatos en las plantillas en ReDIF son indispensables para el funcionamiento del repositorio los cuales facilitaran su organización además serán puntos de acceso para buscar y recuperar información.

## 5. CONCLUSIONES

De acuerdo a lo encontrado sobre la organización de los recursos digitales en RePEc podemos concluir lo siguiente:

- RePEc alberga un total de 2,471,049 registros entre preprints, artículos de libros, capítulos de libros y software, considerándose uno de los repositorios más grandes del mundo.
- Existen dos posibilidades de colaboración en RePEc: 1) proporcionando datos, documentos o archivos y/o 2) ofrecer servicios a los usuarios finales basado en los archivos, es decir, la primera parte, son servicios de alojamiento de metadatos y visualización de la información; mientras la otra parte son servicios derivados del análisis de los archivos y metadatos que tiene la base de RePEc.
- Por falta de presupuesto, la descentralización y la colaboración son una excelente alternativa que deben considerar los repositorios para organizar sus contenidos. En el caso de RePEc es más práctico dejar que cada institución gestione su propia colección a nivel local, descentralizada, además porque cada uno conoce mejor la información que publica y luego hacerlos centralmente accesible en una interfaz común a través del portal de RePEc. Asimismo, la descentralización brinda menor presupuesto en el mantenimiento del sistema de RePEc, aunque puede haber registros de poca calidad en su descripción.
- RePEc maneja grandes cantidades de registros documentales gracias a su estructura, la cual puede resumir en tres partes: Archivos, core site y servicios. No obstante, gran parte de su funcionamiento es gracias al uso de las plantillas de descripción documental con el formato de metadatos ReDIF (Research Documentation Information Format) para la organización del contenido y el protocolo Guildford para el intercambio de datos.
- El uso de los metadatos ReDIF (Research Documentation Information Format) favorecen en la descripción de los documentos como: preprints,

artículos de revistas, libros, capítulos de libros y aplicaciones de software.

- El protocolo Guildford es necesario para regular los procesos comunicativos de publicación e intercambio de metadatos en Internet. Para regular la comunicación entre un archivo y los servicios, naturalmente se requiere un archivo que contenga un conjunto de directorios y subdirectorios accesibles a través de un servidor ftp o http, lógicamente la parte más importante del protocolo es aquella donde se describe tal estructura.
- La pieza clave de la organización de los documentos en el repositorio de RePEc sigue siendo la descripción bibliográfica a través de las plantillas ReDIF, que nos va a permitir una fácil localización y recuperación de la información. Por lo tanto, juega un papel trascendental el tema de las normas utilizadas para la descripción de los documentos. Unas buenas descripciones nos ofrecerán mayores garantías de que localizaremos la información pertinente a nuestras necesidades de información.
- Varios campos de las plantillas ReDIF tienen inscritos directa e indirectamente elementos de Dublín Core.
- A pesar de que RePEc trabaja con sus colaboradores de archivos de forma descentralizada, tiene descripciones óptimas, considerando que los colaboradores llenan plantillas y pueden llegar a cumplimentar de forma incorrecta algún campo, pero sería una tarea muy ambiciosa que el repositorio revisará las descripciones, sin embargo, su sistema tiene integrado un filtro de errores de sintaxis que indica los inconvenientes de tipo técnico más que de contenido.
- La estructura que tiene el repositorio de RePEc puede ser adquirido para crear repositorios institucionales o académicos.

## REFERÊNCIAS

BARRUECO CRUZ, J.M.; KRICHEL, T. Sistemas de información-Acceso a prepublicaciones sobre economía: RePEc. **El Profesional de la Información**, v. 8 n.12, p. 28-35. 1999.

BARRUECO CRUZ, J.M.; KRICHEL, T. Prepublicaciones: distribución centralizada vs. descentralizada. **Revista Española de Documentación Científica**, v.23, n.2, p.137-147. 2000.



BARRUECO CRUZ, J.M.; KRICHEL, T. Distribución de publicaciones a través de Internet el modelo RePEc. **Coloquio de la Asociación Internacional de Bibliología**, n. 15.p.1-9. 2000b.

BARRUECO CRUZ, J.M.; KRICHEL, T. Cataloging economics preprints: an introduction to the RePEc project. **Journal of Internet Cataloging**, v.3, n.2-3, p.227-241. 2000c.

BARRUECO CRUZ, J.M.; KRICHEL, T. Distributed cataloging on the Internet: the RePEc project. In: GREEMBERG, J. (Org.). **Metadata and organizing educational resources on the internet**. New York: Haworth Information Press, 2000d. p. 227-241.

BARRUECO CRUZ, J.M.; KRICHEL, T. (2005). Building an autonomous citation index for GL: RePEc, the Economics working papers case. **Grey Journal**, v.1, n.2, p. 91-97.

EDO HERNÁNDEZ, V. Las clasificaciones de materias en Economía: valoración crítica y propuesta de una alternativa básica general. **Revista Española de Documentación Científica**, v.33, n.4, p. 600-623. 2010.

HILLMANN, D. **Using Dublin Core - The Elements**. 2011. Disponível em: <<http://dublincore.org/documents/usageguide/elements.shtml>>. Acesso em:18 dez. 2017.

KARLSSON, S.; KRICHEL, T. RePEc and S-WoPEc: Internet access to electronic preprints in Economics. In: **IFIP Conference on Electronic Publishing in Ronneby**, May 10-12, 1999. p. 204-214.

KRICHEL, T. About NetEC, with special reference to WoPEc. **Computers in Higher Education Economics Review**, v.11, n. 1, p.19-24. 1997.

KRICHEL, T. (1999). **Guildford protocol**. Disponível em: <<http://ideas.repec.org/p/rpc/rdfdoc/guildp.html>>. Acesso em:18 dez. 2017.

KRICHEL, T. RePEc, an open library for Economics. In: JEFFREY K.; MACKIE-MASON, W.; PRADT, L. (Ed.). **The Economics and usage of digital library collections**, University Michigan, 2008. p. 1-6.

KRICHEL, T. (2007). **ReDIF version 1**. Disponível em: <[http://openlib.org/acmes/root/docu/redif\\_1.html](http://openlib.org/acmes/root/docu/redif_1.html)>. Acesso em:18 dez. 2017.

KRICHEL, T.; ZIMMERMANN, C. The economics of open bibliographic data provision. **Economic Analysis and Policy**, v.39, n.1, p.143-152. 2009.

LÓPEZ-BORRULL, A. Evolución de repositorios temáticos y megarevistas: visión 2017. **Anuario ThinkEPI**, v. 11, p. 242-246. 2017.

PÉREZ MATOS, N.E. De la descripción bibliográfica a la asignación de metadatos: un llamado al orden. **ACIMED: Revista Cubana de los**

**Profesionales de la Información y la Comunicación en Salud**, v.14, n.6, p.1-11. 2006.

RePEc. Research Papers in Economics (RePEc). 2017. Disponível em: <<http://repec.org/>>. Acesso em:18/12/2017.

SEILER, C.; WOHLRABE, K. Ranking economists on the basis of many indicators: An alternative approach using RePEc data. **Journal of Informetrics**, v.6, n.3, p. 389-402. 2012.

SENSO, J.; ROSA PIÑERO, A. El concepto de metadato. Algo más que descripción de recursos electrónicos. **Ciência da Informação**, v. 32, n.2, p.95-106. 2003.

SUBIRATS COLL, I.; BARRUECO CRUZ, J. M. RCLIS: una biblioteca digital distribuída para documentación. In: **III Jornadas de Bibliotecas Digitales:(JBIDI'02): El Escorial (Madrid)**. Grupo de Ingeniería del Software, 18-19 de noviembre, 2002. p. 143-152.

ZIMMERMANN, C. Academic rankings with RePEc Working papers. **Econometrics**, v.1, n.3, p. 249-280. 2013.

## ORGANIZATION OF DIGITAL RESOURCES IN REPEC THROUGH REDIF METADATA

### ABSTRACT

**Introduction:** The disciplinary repository RePEc (Research Papers in Economics) provides access to a wide range of preprints, journal articles, books, book chapters and software about economic and administrative sciences. This repository adds bibliographic records produced by different universities, institutes, editors and authors that work collaboratively following the norms of the documentary organization.

**Objective:** In this paper, mainly, we identify and analyze the functioning of RePEc, which includes the organization of the files, which is characterized using the protocol Guildford and metadata ReDIF (Research Documentation Information Format) templates own for the documentary description. **Methodology:** Part of this research was studied theoretically in the literature; another part was carried out by observing a series of features visible on the RePEc website and in the archives of a journal that collaborates in this repository. **Results:** The repository is a decentralized collaborative project and it also provides several services derived from the metadata analysis.

**Conclusions:** We conclude that the ReDIF templates and the Guildford communication protocol are key elements for organizing records in RePEc, and there is a similarity with the Dublin Core metadata.

**Descriptors:** Repository. Preprinted. desktop publishing. scientific communication. metadata.

## ORGANIZAÇÃO DE RECURSOS DIGITAIS NO REPEC ATRAVÉS DE METADADOS REDIF

### RESUMO

**Introdução:** O repositório disciplinar da RePEc (Research Papers in Economics) fornece acesso a uma ampla gama de preprints, artigos de revistas, livros, capítulos de livros e software sobre temas de ciências econômicas e administrativas. Este repositório adiciona registros bibliográficos produzidos por diferentes universidades, institutos, editores e autores que trabalham de forma colaborativa seguindo as regras da organização documental. **Objetivo:** Neste artigo, identificamos e analisamos principalmente o funcionamento do RePEc, que inclui a organização dos arquivos, que se caracteriza pelo uso do protocolo Guildford e dos modelos de metadados ReDIF (Research Documentation Information Format) para a descrição documental. **Metodologia:** Parte desta pesquisa foi estudada teoricamente na literatura; outra parte foi realizada observando uma série de características visíveis no site RePEc e nos arquivos de um periódico que colabora neste repositório. **Resultados:** O repositório é um projeto colaborativo descentralizado e também fornece vários serviços derivados da análise de metadados. **Conclusão:** Concluímos que os modelos ReDIF e o protocolo de comunicação de Guildford são elementos-chave para a organização de registros no RePEc, e existe uma similaridade com os metadados de Dublin Core.

**Descritores:** Repositório, preprints, publicação eletrônica, comunicação científica, metadados.