

# Acceso a prepublicaciones en Economía RePEc

José Manuel Barrueco & Thomas Krichel

Octubre 1999

## 1 Introducción

La generalización del uso de las redes teleinformáticas está transformando drásticamente los métodos y canales a través de los cuales los científicos e investigadores comunican los resultados de sus trabajos. Esta transformación abarca a los canales formales, con la aparición de las revistas electrónicas, pero, sobretudo, se hace patente en los canales informales. Son estas aquellas formas de comunicación en las que no existe un control de calidad de los contenidos que se transmiten. Es en este campo, y más concretamente en el área de las prepublicaciones, donde los autores han descubierto más pronto las novedades que puede aportar la teleinformática. Unos cambios que podríamos resumir en: mayor libertad para llegar a los lectores. La comunicación a través de redes permite prescindir (hasta cierto punto) de los intermediarios tradicionales en el sistema de publicación de la ciencia, como son por ejemplo los editores y los bibliotecarios.

Con esta idea hemos visto nacer en los últimos años proyectos como:

- **XXX**. Está accesible en la dirección `xxx.lanl.gov`. Fue creado por Paul Ginsparg en 1991, varios años antes del nacimiento del web, para cubrir el campo de la Física de Altas Energías. Actualmente su cobertura temática se extiende no sólo a Física sino también a disciplinas relacionadas como por ejemplo las Matemáticas. Este servicio, con más de 21.000 documentos en texto completo distribuidos de forma gratuita el pasado año, 70.000 consultas por día, más de 35.000 usuarios habituales y mirrors en 16 países constituye un ejemplo excepcional en la distribución electrónica de prepublicaciones.
- **CogPrints**. El impresionante éxito del modelo de Ginsparg ha llevado a su exportación a otras disciplinas. Así tenemos por ejemplo el servidor de Stevan Harnad, CogPrints accesible en la dirección `cogprints.soton.ac.uk/`. Fue creado en 1997 y está financiado por el Electronic Libraries Programm (eLib) en el Reino Unido. Su cobertura temática es muy amplia, abarcando Psicología, Filosofía, Biología, Lingüística, etc.
- **NCSTRL**. Pronunciado "ancestral", es la Networked Computer Science Technical Reference Library. Está accesible en la dirección `www.ncstrl.org`. Al igual que las anteriores, es una colección internacional de informes de investigación y documentos de trabajo en campos relacionados con la Informática.

En el presente trabajo vamos a estudiar una iniciativa semejante aplicada al área de Economía: RePEc. Realizaremos una breve historia del proyecto, una descripción de su modelo de funcionamiento y finalmente evaluaremos su estado actual atendiendo a cuestiones como el número de documentos distribuidos, número de consultas, etc.

## 2 Descripción de RePEc

### 2.1 Qué es RePEc?

RePEc son las siglas de **Research Papers in Economics**. Básicamente RePEc es un conjunto de herramientas conceptuales, protocolos, normas y software cuyo objeto es la distribución electrónica y descripción bibliográfica de los documentos científicos producidos por una disciplina académica, en concreto la Economía

### 2.2 Cómo nació RePEc?

Antes de pasar a describir RePEc es conveniente trazar una breve historia de su evolución que puede ayudar a aclarar los conceptos definidos a continuación.

La arqueología de RePEc se remonta a 1993. Sus orígenes los encontramos en el proyecto WoPEc (Working Papers in Economics) financiado en el Reino Unido por eLib (Electronic Libraries Programm). Este proyecto tenía como objeto inicial la creación de una base de datos de prepublicaciones o documentos de trabajo en formato electrónico y accesibles gratuitamente en internet. Todo ello con el objetivo de facilitar la comunicación entre científicos por medios gratuitos y alternativos a los canales comerciales tradicionales.

El aumento espectacular en el número de este tipo de documentos que se produjo en los años 1995/96 puso de manifiesto la dificultad de elaborar tal base de datos desde una perspectiva centralizada: una única institución recopilando información sobre publicaciones, en la línea de las bibliotecas tradicionales. Así se buscó la colaboración de aquellos departamentos que producían working papers, solicitándoles que remitieran a WoPEc las referencias cada vez que publicaran nuevos documentos. Muy pronto varios departamentos respondieron a esta llamada al constatar las ventajas publicitarias que les reportaba el figurar en una base de datos que por entonces ya era líder mundial en su disciplina.

Pero el problema no acababa aquí. Qué pasaría cuando la financiación de eLib acabara? La magnitud que estaba alcanzando el proyecto hacía imposible su continuidad sin financiación. Dado que el objetivo era permanecer al margen del mercado comercial, proporcionando un producto gratuito, había que buscar soluciones sostenibles de futuro. Dichas soluciones pasaban por la descentralización. Como la propia filosofía de internet, era necesario que el sistema continuara funcionando en el futuro, aunque alguno de sus elementos desapareciera. Para ello se pensó en una descentralización total, en la cual el ideal sería que cada departamento catalogara sus propios documentos para que después esas descripciones bibliográficas se pudieran intercambiar con el resto de departamentos a través de la red. Se generaría así una masa de datos bibliográficos de dominio público que cada departamento o institución podría utilizar en la forma que quisiera. Además no habría ningún **big brother** que controlara el sistema ya que cada departamento sería dueño exclusivamente de sus propios datos.

Con esta idea como telón de fondo, en Mayo de 1997 WoPEc convocó una reunión de colaboradores en Guildford (Reino Unido) a la que asistieron representantes de Suecia y los Países Bajos. Esta reunión es clave, ya que en ella nació el concepto de RePEc y se sentaron las bases del sistema. Allí se perfilaron las dos normas que forman el centro de RePEc, el Protocolo de Guildford el cual establece las normas de colaboración entre las instituciones participantes y ReDIF (Research Documents Information Format) que establece un formato para la descripción bibliográfica de documentos.

Desde entonces RePEc ha crecido de forma espectacular hasta alcanzar actualmente casi cien departamentos colaboradores habituales. Entre ellos están los principales productores de working papers a nivel internacional, el NBER (National Bureau of Economic Research) de Estados Unidos, así como importantes instituciones a nivel mundial en el ámbito de la Economía como son los Bancos Federales de Estados Unidos (FedInPrint) o el CEPR (Center for Economic Policy Research) del Reino Unido.

### 2.3 Cúal es el objetivo de RePEc?

El objetivo básico de RePEc es crear un sistema de intercambio de información bibliográfica y documentos electrónicos entre las instituciones participantes, con el fin de facilitar el acceso por medios electrónicos a los últimos resultados de investigación en las diferentes áreas de la Economía.

Si bien el objeto principal de RePEc son los documentos de trabajo, pues es en esta tipología documental donde primero se recogen los resultados de investigación, con el tiempo ha evolucionado para englobar también artículos aparecidos en revistas y software de distribución gratuita.

Mientras que la distribución de información bibliográfica es gratuita, en el caso de los textos completos de los documentos, los departamentos o instituciones editoras pueden optar entre distribuirlos gratuitamente o cobrar. En el último caso se encuentran, por ejemplo, los artículos publicados en revistas de la editorial Springer. Aquí la información bibliográfica, hasta donde lo permiten las limitaciones de copyright, es libre, pero el acceso al texto completo está restringido a suscriptores del servicio LINK.

### 2.4 Introducción del modelo RePEc

El funcionamiento básico de RePEc se estructura en torno a dos normas, una que establece los principios y reglas de colaboración entre los participantes en el sistema, llamada Protocolo de Guildford y otra, denominada ReDIF, que define el formato para la descripción bibliográfica de los documentos electrónicos. Aunque ambas normas fueron creadas desde la perspectiva de la Economía, son totalmente independientes de la disciplina y pueden ser fácilmente adaptadas a otros campos. Igualmente, aunque el objeto inicial de RePEc es la distribución de documentos, el esquema puede servir también para la distribución de otro tipo de información, como datos sobre investigadores, instituciones, etc.

ReDIF son las siglas de Research Documents Information Format. Este es un formato de datos simple, basado en la estructura "*campo: valor*", e inspirado en los "templates IAFA" (Internet Anonymous FTP Archives). En este sentido, RePEc no sigue un estándar existente o emergente en cuanto a las descripciones bibliográficas, sino que se decidió crear un nuevo formato por razones que veremos más adelante. La ventaja de esta decisión es que el formato ReDIF es mucho más flexible que cualquiera de los existentes y puede ser fácilmente modificado a medida que vayan surgiendo nuevas necesidades. Por otro lado es lo suficientemente simple como para que pueda ser utilizado por personal no especializado, por ejemplo, personal de administración de los departamentos que en muchas ocasiones son los encargados de actualizar las listas de documentos publicados.

El Protocolo de Guildford, cuyo nombre hace referencia a la ciudad donde fue creado, proporciona un conjunto de reglas para la publicación e intercambio de documentos y metadatos en la red. Podría ser implementado por cualquier grupo o individual que deseara intercambiar documentos en internet. Como veremos más adelante establece dos niveles para la participación de los departamentos, el de **archivo**, de carácter pasivo, simplemente proporciona información, y el de **servicio**, activo, ya que toma la información de los anteriores y construye un "servicio" de utilidad para los usuarios finales.

Gráficamente este modelo lo vemos representado en la figura 1

Como vemos en el gráfico, el esquema RePEc está compuesto básicamente por: el **core site**, quien recoge y suministra a los archivos y servicios información básica sobre la localización del resto de participantes. Los archivos, encargados de almacenar la información bibliográfica y en algunos casos los documentos propiamente dichos. Los servicios, que extraen los datos de los archivos y crean una presentación de los mismos para los usuarios finales. Finalmente, los propios usuarios quienes acceden al sistema a través de los servicios con ayuda de un navegador web o un cliente de correo, etc. dependiendo de cada caso.

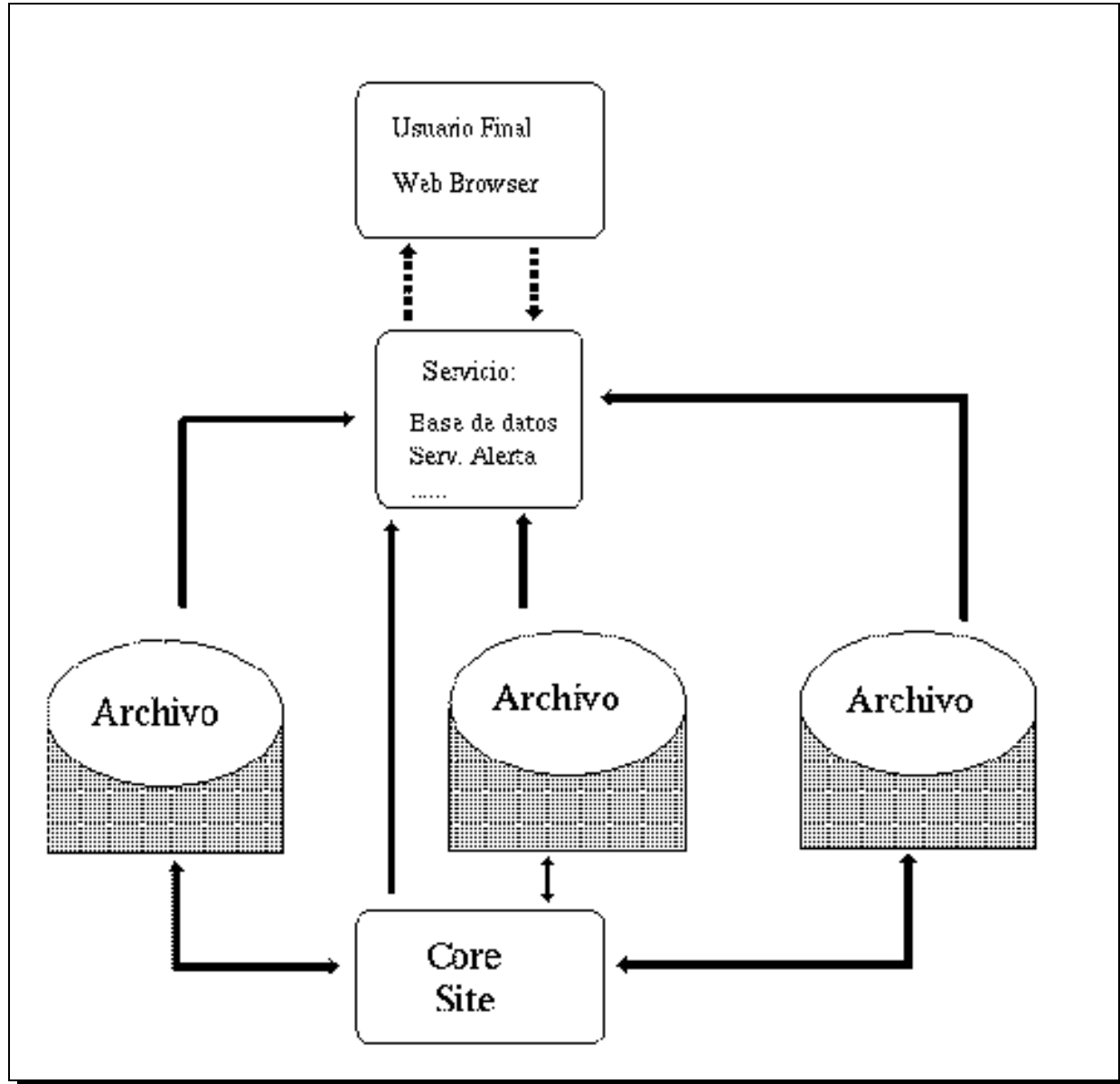


Figure 1: Modelo RePEc

## 2.5 Cuales son los elementos que integran el sistema?

Como ya hemos dicho, las instituciones pueden participar en RePEc de varias formas diferentes, dependiendo de si proporcionan datos o documentos (Archivos) o si ofrecen servicios finales a los usuarios (Servicios).

### 2.5.1 Qué son los archivos?

Archivos son aquellos departamentos o centros de investigación que proporcionan información bibliográfica sobre los documentos que publican. Toda esta información es de dominio público y está accesible gratuitamente en la red. Puede ser copiada y/o distribuida para propósitos de investigación por cualquier persona, relacionada o no con RePEc. Por supuesto puede ser también utilizada con fines comerciales, si el vendedor encuentra clientes dispuestos a pagar por algo que puede conseguir gratuitamente, claro está. El texto completo de los documentos

puede ser también distribuido junto con las referencias si el departamento así lo desea.

Desde un punto de vista técnico, un archivo es simplemente una estructura de directorios y subdirectorios en un servidor FTP o HTTP, donde se almacenan ficheros en formato ASCII conteniendo los datos bibliográficos y en ocasiones ficheros PDF, PS, etc. conteniendo el texto completo de los mismos. Esta estructura debe seguir unas reglas fijas definidas en el Protocolo de Guildford, con objeto de permitir a los robots el acceso a los datos.

Una explicación más detallada de la estructura de un archivo puede verse en el apartado dedicado al Protocolo de Guildford.

**Qué es el core site?** Aunque el objetivo de RePEc es mantener una descentralización máxima, es obvia la necesidad de un cierto control centralizado. Así, de la coordinación del sistema se encarga un archivo central, denominado **core site**. Este archivo está accesible en la dirección `ftp://netec.mcc.ac.uk/pub/NetEc/RePEc/all`. Sus funciones son:

- Mantener actualizada la documentación del sistema. Los dos principales documentos son las definiciones del Protocolo de Guildford y la especificación ReDIF. Además se encarga de mantener las páginas del servidor web de RePEc donde se ofrecen guías con información para la participación en RePEc, documentación del software distribuido por el sistema, etc.
- Regulación de la toma de decisiones entre los miembros del grupo. La comunicación entre los distintos participantes se realiza a través de una lista de discusión: `RePEc-Admin@mailbase.ac.uk`
- Incluir y en su caso retirar archivos del sistema.
- Asignar códigos de identificación a cada nuevo archivo. A cada departamento que desea colaborar con RePEc se le asigna un código único compuesto de tres letras. El asunto de los códigos es clave para el desarrollo de RePEc y sobre ellos volveremos más adelante.
- Distribución de software para el funcionamiento de los archivos y servicios. Básicamente este software incluye programas para controlar la sintaxis de las descripciones bibliográficas, para realizar la copia o mirror de información desde los archivos a los servicios y para crear presentaciones a través del web utilizando las descripciones bibliográficas. Todo el software es mantenido y actualizado por el archivo central, de forma que cualquier departamento participante, aunque carezca de personal informático, pueda crear un archivo o servicio. El lenguaje de programación utilizado es **perl**. Los programas han sido diseñados inicialmente para funcionar en máquinas UNIX, aunque existe una preocupación creciente por crear versiones en otros sistemas operativos como por ejemplo Windows NT.

### 2.5.2 Qué son los servicios?

La información ofrecida por los archivos o departamentos en formato ReDIF es de poca o ninguna utilidad para los usuarios finales tal cual está en los servidores. Así, es necesario que algunos departamentos, que harían las funciones de intermediarios, tomen la información de los archivos y le incorporen un determinado valor añadido para hacerla accesible a los usuarios. Estos son los denominados servicios.

Ese valor añadido puede adoptar distintas formas: la más simple sería convertir las descripciones bibliográficas del formato ReDIF original a un formato legible por una persona, como por ejemplo páginas HTML. También puede consistir en crear un índice de los datos que permita realizar búsquedas, etc. O realizar una selección, comentario y presentación de determinados documentos existentes en el sistema.

En general cualquier archivo puede crear su propio servicio utilizando datos propios y/o procedentes del resto de archivos. Así nos podemos encontrar con departamentos que ofrecen archivos, el caso más frecuente; departamentos que ofrecen servicios, y departamentos que ofrecen ambas cosas a la vez.

En resumen, nos encontraríamos ante una colaboración descentralizada, con multitud de servidores de información, que comparten sus datos con todos los demás servidores, de forma que cualquier miembro del grupo puede hacer uso de los mismos para crear un servicio final al usuario.

### Cuales son los servicios más conocidos?

- **WoPEc** y **BibEc** accesibles en las direcciones  
<http://netec.mcc.ac.uk/WoPEc.html> y <http://netec.mcc.ac.uk/BibEc.html>

son los servicios más antiguos. Son dos bases de datos que ofrecen información sobre documentos de trabajo en formato electrónico e impreso respectivamente. Ambos pueden utilizarse de dos formas: o bien ojeando las paginas html estáticas que contienen las distintas series de los archivos participantes y una clasificación JEL (Journal of Economic Classification) de los mismos, o bien utilizando alguno de los motores de búsqueda ofrecidos. Actualmente se puede elegir entre tres bases de datos: una, la más antigua, utilizando WAIS, otra utilizando un sistema SQL (MySQL) y una tercera a través del software ROADS. En cuanto a número de documentos, WoPEc ofrece información sobre más de 14.000 working papers en texto completo y BibEc proporciona datos bibliográficos sobre más de 45.000 documentos impresos, con información detallada sobre las formas de obtener copias de los mismos.

- **NEP** son las siglas de New Electronic Papers. Es un servicio de alerta informativa a través de listas de distribución, via correo electrónico, donde se anuncian semanalmente las novedades aparecidas en RePEC. NEP se compone de más de 30 listas de novedades temáticas, especializadas en áreas concretas de la Economía, como por ejemplo:

C+latexcommand(ó)digo	T+latexcommand(í)tulo	JEL	Editor
nep-cfn	Corporate Finance		Evgenia Stupina
nep-cmp	Computational Economics		Nienke Oomes
nep-dcm	Discrete Choice Models	c25-c35	Josep M Camacho Cabisco
nep-dge	Dynamic General Equilibrium		Christian Zimmermann
nep-ecm	Econometrics		Sune Karlsson
nep-ets	Econometrics-time series		Yong Yin
nep-eff	Efficiency and Productivity		Vania Sena
nep-env	Environmental Economics		Francisco S. Ramos
nep-eec	European Economics		Marco Catenaro
nep-evo	Evolutionary Economics		Murat Yildizoglu
nep-exp	Experimental Economics		Reinhard Sippel

Cada una de ellas está moderada por un investigador de prestigio en el tema tratado. El funcionamiento de NEP es el siguiente: semanalmente se extraen las novedades del sistema y se envían a los moderadores de cada lista. Por el momento las novedades sólo incluyen working papers cuyo texto completo esté accesible electrónicamente. El resto de tipos documentales son excluidos. Los editores se encargan de seleccionar los documentos que se ajustan al objeto de su lista y distribuirlos entre los suscriptores. Para más información sobre NEP pueden consultarse las páginas del proyecto en la dirección <http://netec.mcc.ac.uk/NEP>.

## 2.6 Cómo se regulan las relaciones entre archivos y servicios?

Como ya dijimos anteriormente, a través de dos normas:

### 2.6.1 El Protocolo de Guildford

Este protocolo establece una serie de normas y recomendaciones que deben seguir todos los archivos que forman parte del sistema. El no seguimiento de las mismas haría imposible la comunicación entre las distintas máquinas.

Si técnicamente definimos un archivo como una estructura de directorios y subdirectorios accesibles en un servidor FTP o HTTP, lógicamente la parte más importante del protocolo es aquella donde se describe tal estructura. A grandes rasgos el protocolo establece que: todo archivo debe ser identificado por un código único (`archivo_id`) en el sistema. Este código es asignado por el core site y estará compuesto por tres letras. Constituye el elemento raíz de la estructura. Los archivos a su vez identificarán las distintas colecciones o series de documentos que publiquen con un código de seis letras (`series_id`). Cada serie dispondrá de un subdirectorio donde se almacenaran las descripciones bibliográficas. Dentro del archivo todos los ficheros con extensión `.rdf` contendrán datos en formato ReDIF.

Veamos un ejemplo. El National Bureau of Economic Research (NBER) proporciona un archivo RePEC, con código `nbr`, en la dirección `http://nberws.nber.org/RePEC/nbr`. En este archivo se ofrece información sobre la colección "NBER working papers", código `nberwo` en el subdirectorio `http://nberws.nber.org/RePEC/nbr/nberwo`. Este subdirectorio contiene una serie de ficheros ASCII con datos bibliográficos sobre los documentos.

### 2.6.2 ReDIF

La primera pregunta que se nos presenta al hablar de ReDIF es por qué un nuevo formato de descripción de documentos? por qué no utilizar alguno de los ya existentes como por ejemplo el MARC o el Dublin Core?

Existen varias razones. En primer lugar la sencillez. Normalmente las descripciones bibliográficas van a ser creadas por personal no especializado, personal de administración, por ello debe ser un formato lo más fácil de utilizar posible. Con ello descartaríamos formatos como el MARC, difícil de entender incluso por los bibliotecarios.

En segundo lugar, debe ser un formato interdisciplinar, que no sea "propiedad" de un determinado grupo de usuarios, como es el caso del BibTeX entre la comunidad de físicos y matemáticos.

En tercer lugar, debe ser un formato fácilmente actualizable para adaptarlo a nuevas necesidades. Debe ser algo dinámico, donde no haya que esperar la decisión de varios comités para incluir una modificación.

Finalmente existen también razones históricas. Como ya hemos dicho RePEC surgió a partir de WoPEC. WoPEC desde el año 1995 venía utilizando el formato IAFA (Internet Anonymous FTP Archive) de la IETF (Internet Engineering Task Force). Este formato pensado inicialmente para la descripción de recursos accesibles a través de servidores ftp contó con un amplio apoyo de la comunidad internet durante los años 1995 a 1997, cuando comenzó a cobrar fuerza el Dublin Core.

Al no encontrar entre los formatos de descripción de recursos existentes alguno que se adaptara a las necesidades mencionadas se decidió construir uno propio que siguiera las líneas básicas del formato IAFA. De hecho ReDIF es una evolución o interpretación del IAFA y en gran parte compatible con él. Como él se basa en una sintaxis sumamente sencilla de: "**campo: contenido del campo**".

También igual que en el caso IAFA, ReDIF ofrece distintos tipos de registros dependiendo del documento que se esté describiendo como podemos comprobar en la figura 2

<i>Nombre</i>	<i>Versión</i>	<i>Usado para</i>
ReDIF-Paper	1.0	working papers
ReDIF-Article	1.0	artículos en una revista
ReDIF-Chapter	1.0	capítulos de libros
ReDIF-Series	1.0	colecciones de documentos
ReDIF-Archive	1.0	archivos
ReDIF-Mirror	1.0	mirrors de archivos

Figure 2: Tipos de registros en ReDIF

En un mismo fichero pueden almacenarse varios registros del mismo tipo. Cada registro debe ocupar un párrafo de texto, es decir, no están permitidas líneas en blanco dentro de los registros. Deben comenzar con un campo que es

obligatorio: **Template-Type:** cuyo contenido es alguno de los tipos enunciados más arriba.

En el resto de los campos no importa el orden en que aparezcan. Ahora bien, ReDIF utiliza el concepto de **clusters**, por lo que los datos referidos a un mismo objeto deben aparecer juntos en la descripción.

Un cluster lo podemos explicar de la siguiente forma: existen una serie de datos que son comunes en varios casos, por ejemplo, cada persona mencionada en un registro puede tener un nombre, un número de teléfono, una dirección postal, etc. De igual forma, cada organización mencionada tendrá un nombre, un número de teléfono, una dirección postal, etc. A ese conjunto de elementos que son comunes a la persona y a la organización es a lo que llamamos cluster. Así un cluster sería un conjunto de campos enlazados lógicamente en un registro y que se refieren a un objeto determinado. Todos esos campos deben aparecer seguidos dentro del registro.

ReDIF diferencia tres clusters: "person", "organization" y "file". Seguidamente se presenta un ejemplo del cluster persona:

### Persona

```
Author-Name: Joe Smith
Author-Email: JoeSmith@some.uni.edu
Author-Postal: PO Box 123, Smith Street, Somewhere In The, Universe, 987654
Author-Phone: +99 456-321123
Author-Homepage: http://www.some.where.edu/~JoeSmith
```

El registro se completa además con un conjunto de campos simples como son por ejemplo, title, abstract, creation-date, classification, keywords, handle.

Mención especial merece este último ya que es un campo obligatorio cuyo contenido es un código que identifica unívocamente a cada objeto descrito en RePEC. Esta es la característica más importante del formato ReDIF, ya que una simple cadena de caracteres nos proporciona información sobre el archivo, la serie y el número de un objeto determinado. Por otro lado a través del handle podemos crear una estructura relacional que enlace los distintos elementos que forman la descripción de un documento. El handle se compone de: la cadena RePEC seguida de " : ", el código del archivo al que pertenece el objeto, " : ", el código de la serie donde se ha publicado, " : ", un número de orden asignado por el archivo. Así tendríamos:

```
Handle: RePEC:bon:bonnsf:a452
Handle: RePEC:die:calsdi:9338
```

Finalmente, un registro ReDIF completo quedaría como sigue:

```
Template-Type: ReDIF-Paper 1.0
Title: A Discussion of the Reliability of Results Obtained with Long-Run Identifying Restrictions
Author-Name: St-Amant, P.
Author-Email: St-Amant@uv.es
Keywords: CENTRAL BANKS ; MONETARY POLICY ; MACROECONOMICS
Length: 14 pages
Classification-JEL: C1 ; E27
Creation-Date: 1998
File-URL: http://www.nber.org/papers/w6490.pdf
```



```

File-Format: application/pdf
File-Restriction: Access to the full text is restricted. Look
up http://www.nber.org/wwpfaq.html for details, or write to
feenberg@nber.org. If you have no access to the full text you will be
shown an abstract page instead. Anyone browsing the
NBER working paper database from a site with a TLD in a non-OECD, non-OPEC
country will be offered full text downloads for any paper for which the
full text is available online, but only if their DNS does reverse name
lookup.
Handle: RePEc:nbr:nberwo:6490
Price: $5.00 per paper (plus $10.00 postage for orders outside U.S.)

```

Este registro ilustra cómo la colección no está necesariamente limitada a documentos gratuitos. Dado que las condiciones de acceso a estos documentos de pago varían el texto del campo File-Restriction es libre.

### 3 Evaluación

En esta sección estudiaremos a modo de evaluación algunos indicadores que nos permitirán conocer el "estado de salud" de RePEc, tanto en lo referido a los archivos como a los servicios.

#### 3.1 Archivos

El primer aspecto a estudiar de los archivos en RePEc es su número. Desde su nacimiento en mayo de 1997 hasta la fecha el número de archivos se ha ido incrementando de forma constante hasta llegar a los más de cien en la actualidad. La evolución durante el pasado año, de abril de 1998 a Agosto de 1999 concretamente, la podemos seguir en la figura 3.

Si nos centramos en los documentos a texto completo distribuidos por los archivos, en la figura ?? tenemos la evolución del número de los mismos por tipo documental: documentos de trabajo (Pre) y artículos publicados en revistas (Art). Aquí podemos comprobar claramente como el objeto de RePEc son las prepublicaciones, mientras que la presencia de artículos es casi testimonial. En ambos casos el número de documentos mantiene una tendencia creciente constante, pero si bien el número de artículos se ha mantenido prácticamente invariable, en torno a los 1.500, durante el periodo analizado; el de working papers ha crecido de forma sustancial, pasando de los 7.500 iniciales a los casi 14.500 de la actualidad. De continuar esta tendencia podemos prever que el número total se duplicará en menos de un año, lo que nos demuestra una vez más el auge que está experimentando esta tipología documental.

La distribución de documentos por archivo, sin distinguir texto completo de información referencial, la podemos ver en la figura 5. Por motivos de espacio solamente incluimos los diez archivos más grandes en RePEc. Destaca a primera vista el tamaño del archivo **fth**, que supera en más del doble de documentos al segundo de la lista. Este archivo está elaborado por la biblioteca del Departamento de Economía de la Universidad de Montreal en Canada, que alberga una de las colecciones de working papers impresos más grandes del mundo. Recoge los fondos de dicha biblioteca y contiene exclusivamente información referencial.

Le siguen en importancia **fip** y **nbr**. Estos son los archivos de los Bancos Federales de los USA y del National Bureau of Economic Research (USA) respectivamente. Como ya hemos mencionado estos son los dos mayores productores de preprints a nivel mundial. A diferencia del primero, estos archivos contienen exclusivamente información de documentos, tanto artículos como preprints, publicados por las instituciones a las que pertenecen. Gran parte de ellos los encontraremos a texto completo.

En cuarto lugar nos encontramos con **wop**. Este es el archivo que podemos considerar el origen de RePEc pues

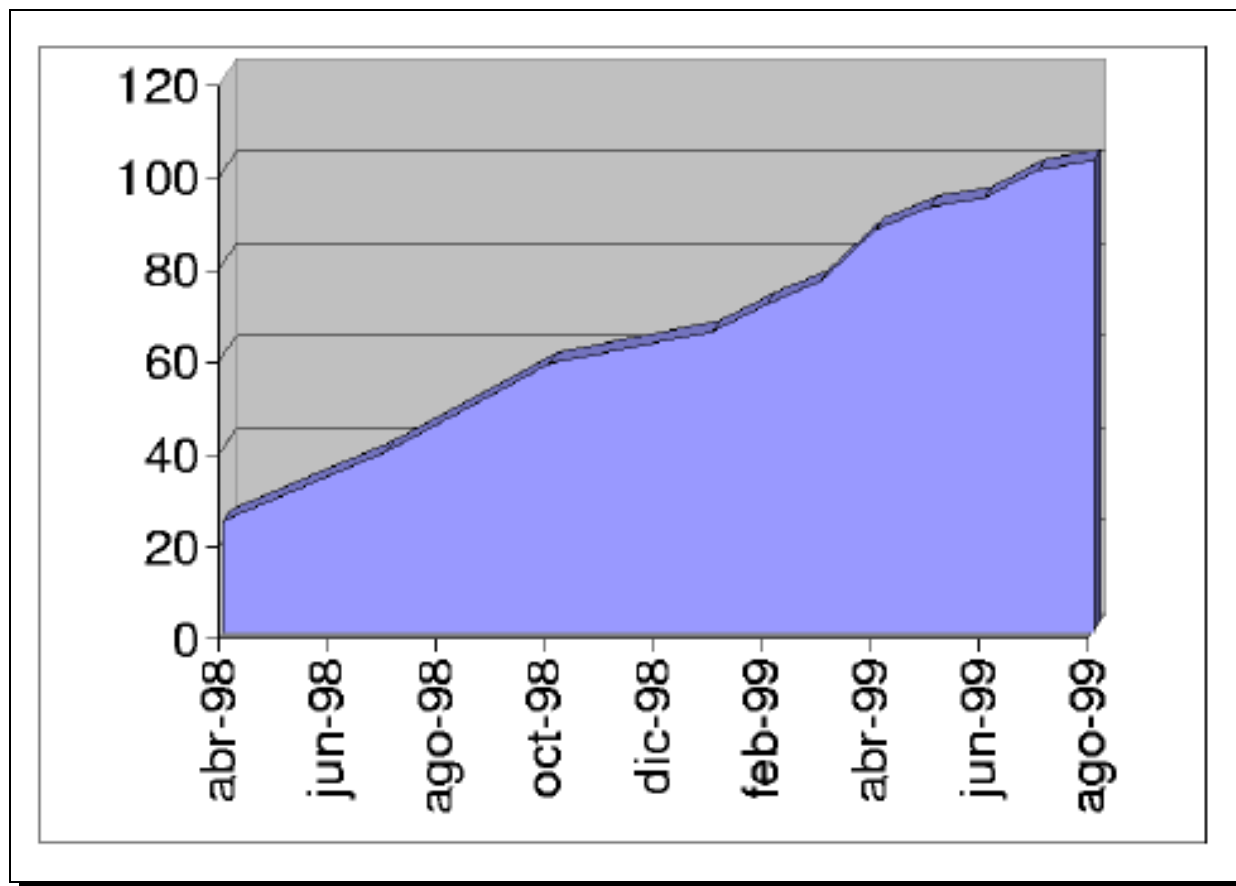


Figure 3: Numero de archivos en RePEc

contiene los fondos recopilados por el antiguo proyecto WoPEc. Son documentos de trabajo cuyo texto completo está disponible en forma electrónica a través de internet.

Del resto de archivos cabe destacar **wpa**, el archivo del Economics Working Papers Archive. Este proyecto fue creado por Bob Parks en 1993 en la Washington University at Saint Louis (USA), con el objetivo de exportar a la Economía el modelo de distribución centralizada de prepublicaciones que ya triunfaba en Física. Durante mucho tiempo fue un competidor directo de WoPEc al intentar imponer un modelo diferente. En la actualidad su objetivo no se ha cumplido y pasa por momentos difíciles debido a problemas de infraestructura y del limitado número de documentos que recibe.

También tenemos **dgr**, otro de los fundadores de RePEc. Este archivo creado por el proyecto DEGREE en los Países Bajos, recoge el texto completo de todos los documentos de trabajo publicados por las universidades y centros de investigación de dichos países.

El resto en su mayor parte son pequeNYos archivos que simplemente distribuyen los documentos publicados por un Departamento. La media de documentos por archivo se sitúa entorno a los cincuenta.

Finalmente vemos en la figura 6 la distribución de los archivos por países de procedencia. El lugar más destacado lo ocupa el Reino Unido con 33 archivos. Esta predominancia es lógica si tenemos en cuenta que es en este país donde nació RePEc y de donde procede la única financiación de la que dispone hasta el momento. Una financiación que incluía la condición de potenciar el uso y participación de departamentos británicos en el proyecto. Le siguen USA y Canadá con 21 y 13 archivos respectivamente, Alemania con ocho y España con cinco: Universidad de Valencia, Universidad Pompeu Fabra, Universidad Complutense de Madrid, Universidad Pública de Navarra y FEDEA.

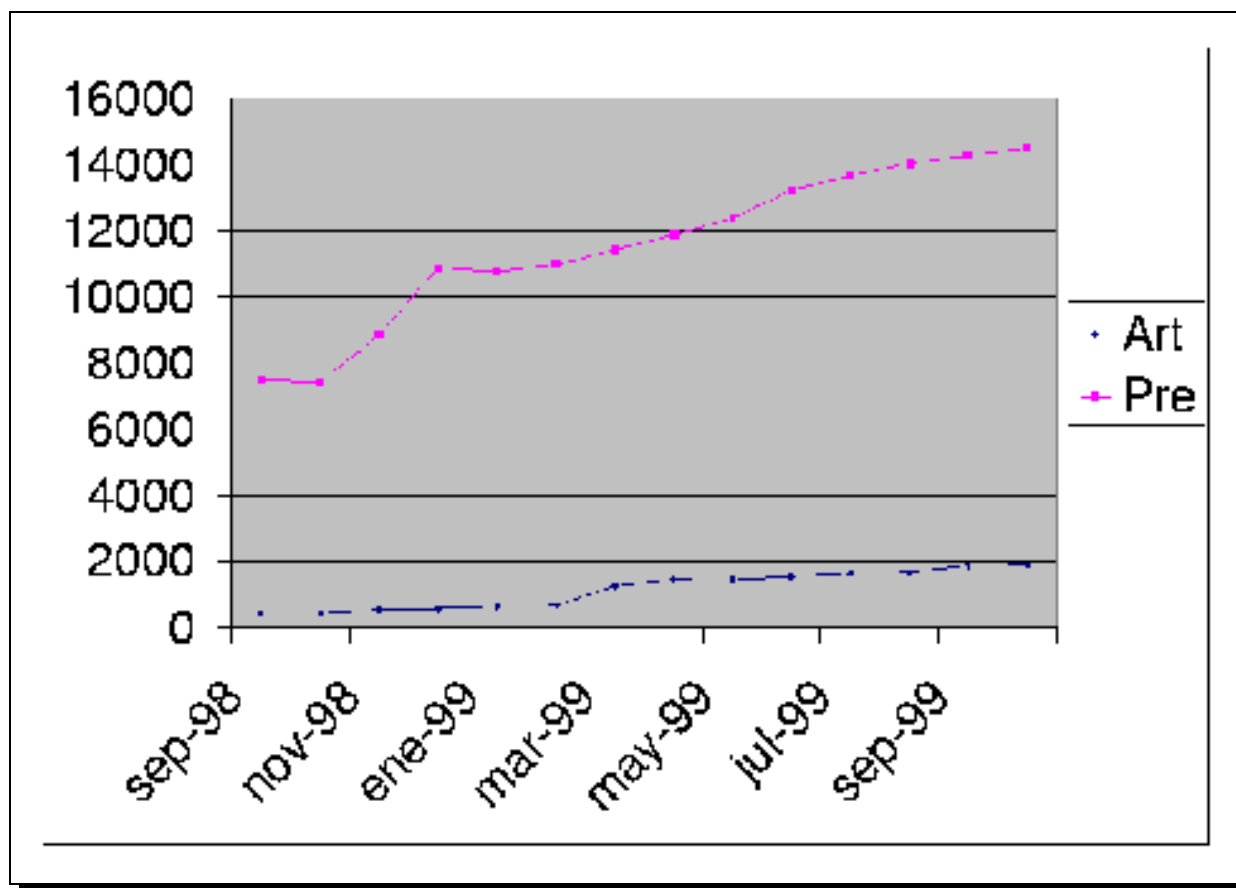


Figure 4: Numero de documentos electronicos en RePEc

Para finalizar el análisis de los archivos podemos concluir que: el número de instituciones interesadas en unirse a RePEc mantiene un crecimiento constante. Unas instituciones estas que proceden principalmente del Reino Unido, USA y Canada. La misma tendencia podemos observar en el número de documentos. En este aspecto, el objeto de RePEc son los documentos de trabajo, mientras que los artículos de revistas mantienen una presencia testimonial.

### 3.2 Servicios (WoPEc)

Desde el punto de vista de los diferentes servicios existentes hemos elegido WoPEc para nuestra evaluación. Recordamos que WoPEc recoge exclusivamente documentos de trabajo y artículos de revista cuyo texto completo esté disponible electrónicamente a través de internet. Vamos a estudiar aspectos de este servicio como la evolución del número de consultas realizadas al sistema y número de documentos descargados del mismo.

En la figura 7 podemos apreciar la evolución del uso de WoPEc en los últimos dos años y medio (desde Enero de 1997 a Junio de 1999). Como vemos el crecimiento en el número de consultas ha sido casi constante, exceptuando dos pequeños baches a finales de 1997 y mediados de 1998. Este número ha pasado de unas 75.000 páginas descargadas mensualmente a comienzos de 1997 a más de 350.000 descargadas el pasado mes de Abril, es decir, se ha multiplicado por cuatro el número de consultas. En el ámbito de este trabajo entendemos por consulta el número de páginas html descargadas del servidor. Es decir, un usuario en la misma sesión puede descargar cuatro, cinco o veinte páginas.

Logicamente estos datos no ofrecen una visión objetiva del servicio. Haciendo un simil con la biblioteca tradicional, si bien es interesante saber cuanta gente nos visita, o al menos viene a ver nuestras instalaciones, mucho más lo es saber cuanta de esa gente hace un uso efectivo de la biblioteca, utiliza el servicio de prestamo o de refe-

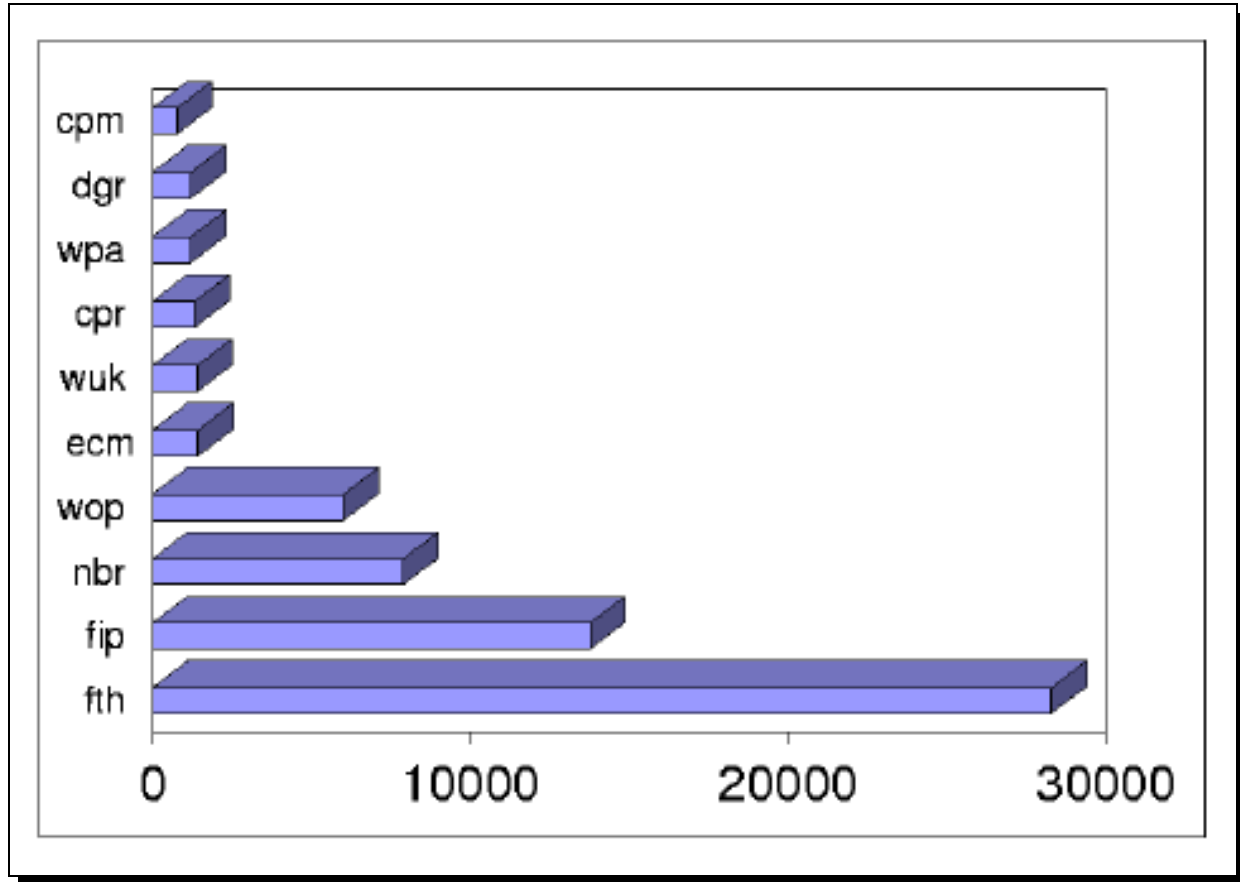


Figure 5: Archivos con mayor número de documentos

rencia, etc. En el caso de WOPeC el aspecto más importante a estudiar es el número de documentos descargados, es decir, aquellos que el usuario además de leer la descripción bibliográfica, se ha tomado la molestia de descargar su texto completo. Un paso adelante más para medir la calidad de los contenidos de WOPeC sería analizar cuantos de ellos de verdad han respondido a las necesidades del investigador, pero eso queda fuera del alcance de este trabajo.

Para estudiar los documentos descargados de la base de datos hemos analizado las estadísticas del servidor durante un periodo de 30 días, del 1 al 30 de Septiembre de 1999. Durante este periodo se descargaron del servidor un total de 35.396 documentos, con una media diaria de 1.179 documentos. 12.899 máquinas diferentes han usado el servicio, lo que nos da una media de 2.7 documentos descargados por cada máquina. En la figura 8 relacionamos las máquinas que han descargado más de 100 documentos, los mayores usuarios de WOPeC:

## 4 Conclusiones

A lo largo de este trabajo hemos descrito y evaluado el sistema de distribución por medios electrónicos de datos, tanto bibliográficos como documentales, que está siendo utilizado actualmente en Economía.

Como conclusión del mismo podemos resaltar en primer lugar la novedad que representa RePEc en relación con otros proyectos similares. Una novedad que se refiere no sólo al máximo grado de descentralización de todas las actividades, sino también al objeto mismo del propio proyecto. Mientras el objeto de bases de datos como XXX es el documento mismo, tratan de almacenar y distribuir documentos, en RePEc el objeto de intercambio son

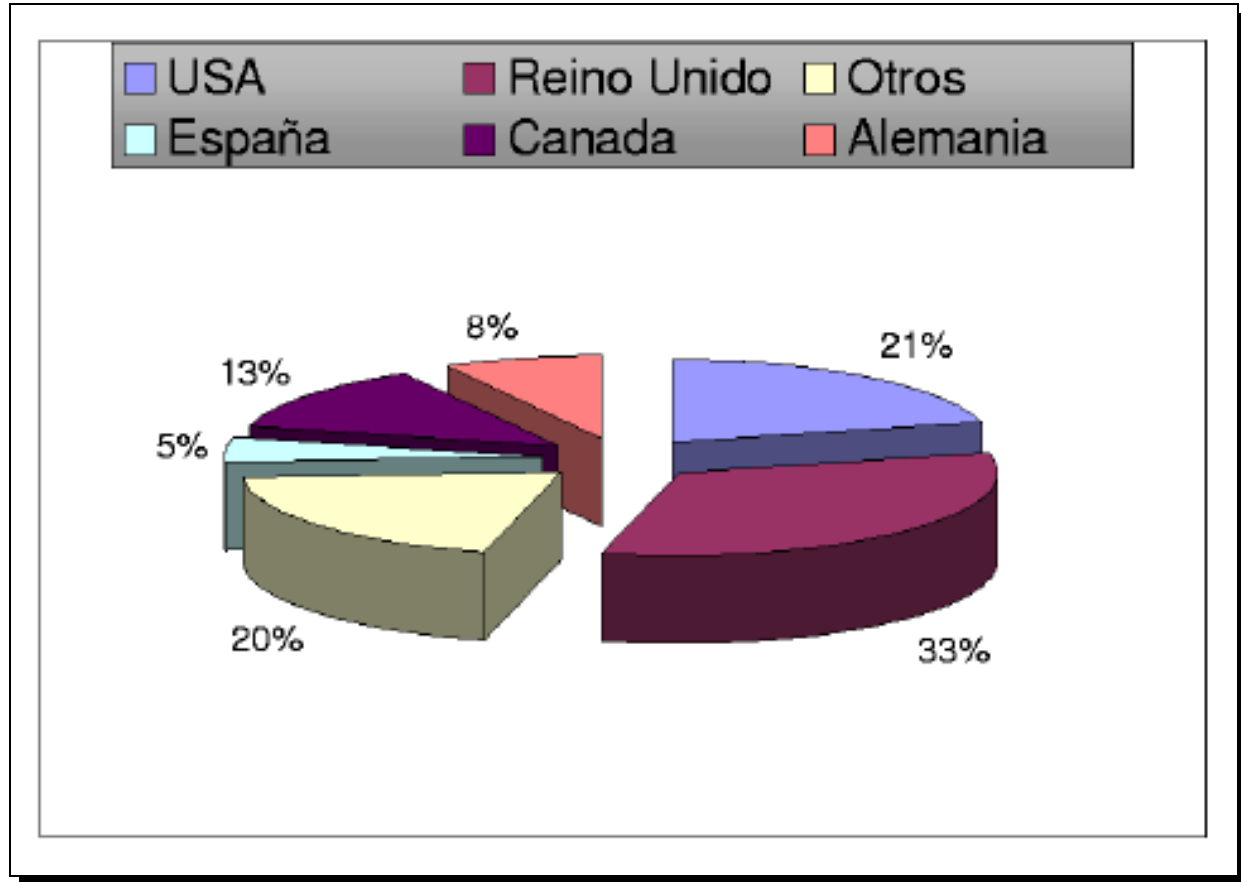


Figure 6: Distribucion de archivos por paises

las descripciones bibliográficas de esos documentos. El texto completo de los mismos puede permanecer en las instituciones que los han generado. En definitiva, el autor sigue manteniendo la propiedad de su obra, no se ve obligado a remitirla a terceras partes. Esta es una diferencia de gran importancia cuando tratamos cuestiones como los derechos de copia, que quedan fuera del alcance de este trabajo.

Por otro lado y a pesar de lo que pueda desprenderse de la documentación, RePEc es un sistema muy fácil de implementar y sobretodo de mantener. Un indicador que nos permite confirmar esta afirmación es el elevado número de departamentos que se están sumando al sistema.

Finalmente, el número creciente de consultas que reciben los servicios, como hemos demostrado con WoPEc, el cual ha triplicado la cantidad de accesos en dos años, augura un buen futuro para RePEc. La buena acogida que está teniendo entre los usuarios se verá incrementada a corto plazo con una inclusión masiva de artículos de revistas y a medio plazo con la inclusión de nuevas tipologías documentales como las series estadísticas, directorios de economistas, etc.

## 5 Bibliografía

Guildford Protocol. RePEc : London, 1997. Accesible en:  
<http://netec.mcc.ac.uk/RePEc/GuilP.html>

Lim, David. Preprint servers: a new model for scholarly publishing?. Australian Academic and Research Libraries (AARL) 27, 1, March 1996. 21-30.

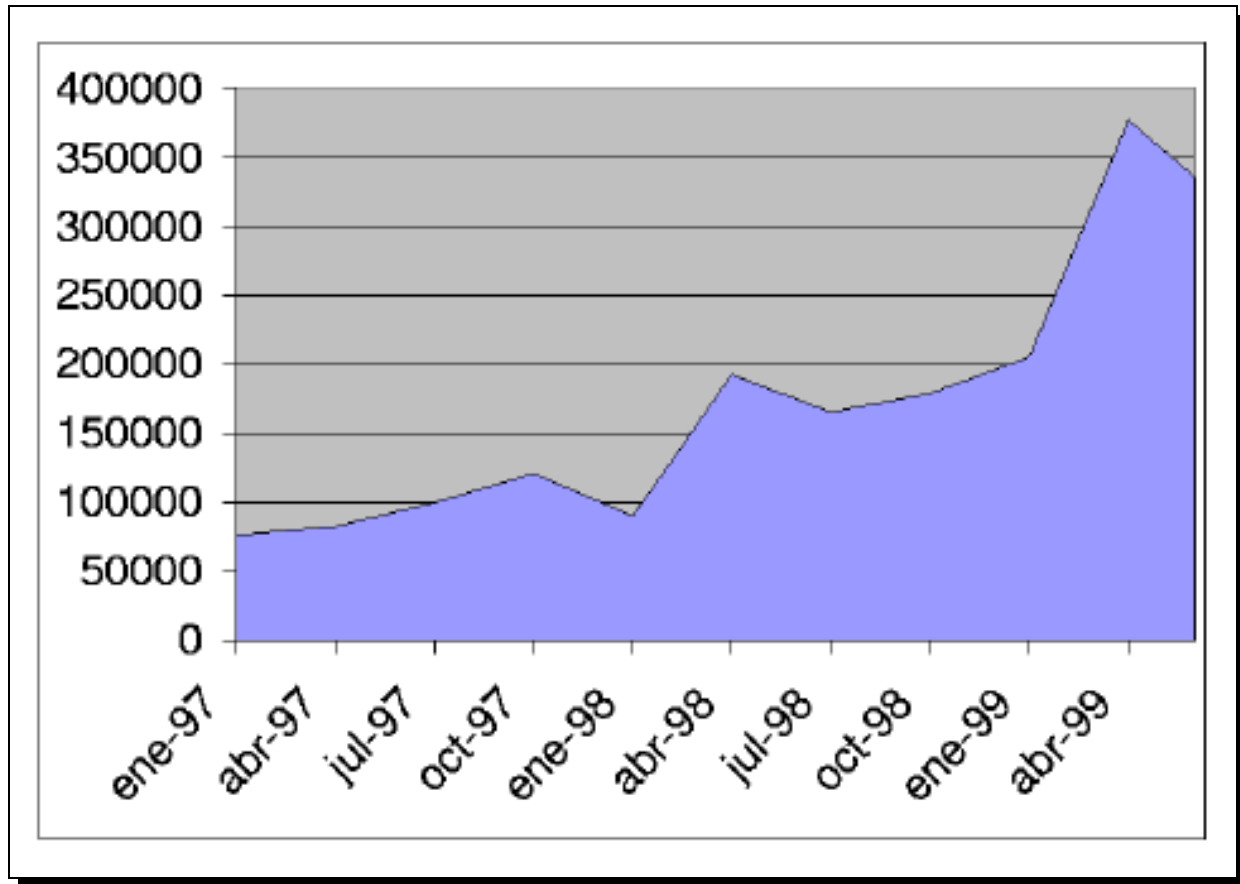


Figure 7: Evolucion de las consultas a WoPEc

Máquina	No. IP	Descargas
pump1.york.ac.uk	144.32.128.3	196
news.nida.ac.th	203.154.126.3	190
lumen.brunel.ac.uk	134.83.176.45	185
	202.28.18.18	146
suws2.rdg.ac.uk	134.225.16.2	124
plum2.bncc.edu.cn	202.205.241.2	124
	147.47.1.102	114
	202.103.164.22	111
	133.5.140.201	109
zsolti.bakats.tvnet.hu	195.38.106.14	107
cache1.jaring.my	161.142.78.71	104
ceres.brunel.ac.uk	134.83.108.8	103
cache41-noc.kornet21.net	168.126.72.13	102

Figure 8: Mayores usuarios de WoPEc

Research Documents Information Format (ReDIF). RePEc : London, 1997. Accesible en: <http://netec.mcc.ac.uk/RePEc/ReDIF.html>

Taubes, Gary. Electronic preprints point the way to "author empowerment". Science, vol. 271. no. 5250, 9 de Febrero de 1996, p. 767.

Youngen, Gregory K. Citation patterns of traditional and electronic preprints in the published literature. College

and Research Libraries, August 1998.

Halpern, Joseph Y. ; Lagoze, Carl. The Computing Research Repository: promoting the rapid dissemination and archiving of computer science research. Documento electrónico disponible en: <http://www.ncstrl.org>