

## MÓDULO 2

### Almacenes de Recursos Abiertos: Repositorios

Por *Pablo García Sánchez*

Profesor del Departamento de Arquitectura y Tecnología de los Computadores de la UGR

---

#### 1. ¿QUÉ ES UN REPOSITORIO?

En el mundo del Software Libre es muy conocido y utilizado el concepto de repositorio. Si bien los más conocidos son los repositorios de software, o de código fuente, existen muchos otros tipos de repositorios, la mayoría de los cuales ofrecen una licencia donde los creadores ceden algunos derechos para permitir usar su contenido libremente.

Un repositorio no es más que un lugar donde almacenar información de algún tipo, que puede ser útil para cierto tipo de personas y sobre la que se mantiene algún tipo de inter-operatividad e histórico sobre los cambios, bajas o altas de dicha información.

El objetivo principal de cualquier repositorio es recopilar, gestionar, difundir y reservar la información que contenga manteniendo en la mayoría de los casos esta información con criterios de ordenación conocidos y en un caso ideal, debería ser interoperable garantizando así el impacto y la visibilidad de la misma. Por último, pero no menos importante, no solo se almacena la información tal cual, sino también la llamada metainformación o los metadatos de los recursos que se hayan incluido en el repositorio como por ejemplo la autoría del mismo, las fechas de publicación, etc. Los metadatos se definen como “datos sobre los datos” y sirven para describir un contenido, llamado recurso. Por ejemplo, un libro en una biblioteca sería un recurso, y sus metadatos serían el nombre del autor, el ISBN, el nombre de la editorial o el código donde está almacenado para facilitar su búsqueda.

Como se ha dicho anteriormente, la mayoría de repositorios cuentan con una licencia indicando que se puede hacer y que no con el contenido puesto a disposición. Como se busca la difusión, la mayoría de los repositorios cuentan con alguna licencia libre, como por ejemplo *Creative Commons* o la *GNU General Public License*.

## 2. REPOSITARIOS DIGITALES

Dependiendo de cuál sea el lugar donde se almacene o el tipo de información que se almacene, tendremos unos tipos de repositorios u otros.

Por ejemplo, y aplicando la definición tal cual, una biblioteca es un repositorio donde el tipo de información que se guarda es los documentos o libros que contiene, así como la ubicación de los mismos dentro del espacio, es decir, qué estanterías, estancias o archivos están ocupando estos documentos de forma que puedan ser encontrados por cualquier persona que quiera acceder a ellos.

Pero los más importantes hoy en día no son las bibliotecas de libros o documentos, sino los llamados **repositorios digitales**, en los que por supuesto se incluyen ya muchas bibliotecas del mismo tipo. En este caso se atiende a una clasificación que tiene en cuenta el soporte donde se almacena la información estableciendo repositorios digitales y no digitales, aunque estos últimos, hoy en día son mucho menos utilizados.

En el caso de los repositorios digitales, los recursos que lo forman son de este mismo tipo y el sitio donde se almacenan no es más que una base de datos en un espacio digital en la que se almacena el recurso, junto con los llamados metadatos (definidos anteriormente), ofreciendo siempre una forma de interoperar con dicha información para mantener siempre un histórico de que ha ocurrido con ella desde que se añadió al espacio digital que lo almacena. Este histórico puede almacenar información ya no solo del recurso, sino de por ejemplo quien lo ha utilizado, en qué fechas, si se modificó de alguna forma, etc. Estos metadatos facilitan que puedan indexarse más rápidamente y ser encontrados por internet con más facilidad de forma automática. Además, muchos de ellos ofrecen APIs (Application Layer Interfaces) que permiten obtener los metadatos desde cualquier lenguaje de programación a través de Internet de una forma sencilla y sin tener que conocer por ejemplo la estructura que los está almacenando.

De hecho, muchos de los repositorios actuales están amparados bajo la iniciativa Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH) [6], que tiene como misión desarrollar y promover estándares de interoperabilidad para facilitar la difusión eficiente de contenidos en Internet. Actualmente cuenta con más de 3700

repositorios afiliados, que pueden consultarse en <https://www.openarchives.org/Register/BrowseSites>. Entre ellos está el Repositorio Institucional de la Universidad de Granada (<https://digibug.ugr.es>) o el Portal de revistas de la UGR (<https://revistaseug.ugr.es/>).

### 3. TIPOS DE REPOSITORIOS DIGITALES

Centrémonos ahora en las clasificaciones que podemos encontrar de repositorios digitales. Aunque existen varios criterios de clasificación, la clasificación presentada en este documento no es estricta, ya que existen repositorios que pueden ser enmarcados en más de un apartado, por ejemplo, un repositorio de datos económicos institucionales podría clasificarse como un repositorio institucional, a la vez que datos y si todos los datos son económicos, podría tratarse como un repositorio temático sobre economía. En cada clasificación mostraremos algunos ejemplos, algunos de los cuales serían desarrollados en más profundidad más tarde en este documento.

#### 3.1 POR EL TIPO DE DATOS ALMACENADOS

En esta clasificación se dividen los repositorios con respecto a una determinada naturaleza, o tipo de información a almacenar: por ejemplo, repositorios de imágenes, repositorios de aplicaciones o herramientas digitales, o los más extendidos: los repositorios de documentos. En esta clasificación no nos interesa la temática de los documentos almacenados. Por ejemplo, en un repositorio de imágenes pueden almacenarse imágenes médicas o imágenes de paisajes, y lo que tienen en común es el tipo de recurso que lo compone, imágenes, y no de qué tratan dichas imágenes.

##### 3.1.1. DOCUMENTOS

Es el tipo más común de repositorio. Normalmente los documentos almacenados suelen ser ficheros en formato PDF junto con cierto tipo de metadatos asociados: autores, versión del documento, fecha de edición, palabras clave para facilitar la búsqueda, etc. Los más extendidos son los repositorios de documentos científicos.

Algunos ejemplos son Arxiv o PubMed, pero también existen repositorios institucionales (que definiremos más adelante) de documentos, como DialNet.

## 3.1.2. DATOS/DATASETS

En estos repositorios se almacenan como recursos conjuntos de datos. Por ejemplo, la temperatura de una ciudad a lo largo del tiempo o el número de habitantes censados a lo largo del tiempo. Normalmente los datos almacenados siguen algún formato abierto (como por ejemplo CSV (Comma-Separated Values: columnas de datos separados por un carácter especial reconocible fácilmente, como es un tabulador), JSON (permite almacenar información en modo texto con cierto orden) [10], o incluso formatos de hojas de cálculo como Libre Office o Excel (este último es un formato propietario). Cada conjunto de datos (o datasets) tiene asociados sus metadatos, como hemos indicado anteriormente.

Un ejemplo muy importante son los repositorios de datos científicos. En ellos los investigadores de todo el mundo comparten los resultados de sus experimentos para que otros científicos puedan reproducir sus investigaciones. En [4] se describen algunos de estos tipos de repositorios. Un ejemplo es Biogrid [3], que es una base de datos biológica de interacciones proteína-proteína, interacciones genéticas, interacciones químicas y similares. Igualmente, debido a que se busca favorecer la difusión y el uso de estos datos, estos recursos también suelen contar con alguna licencia de uso libre.

## 3.1.3 SOFTWARE

Un repositorio de software es un lugar de almacenamiento de paquetes de software [11]. A menudo se almacena una tabla de contenidos, así como sus metadatos asociados. Estos repositorios agrupan datos llamados paquetes. A veces estos paquetes sirven para ser usados por personal más técnico, (por ejemplo, el repositorio Maven ofrece paquetes para el lenguaje de programación Java), y otras veces ofrecen software para un sistema operativo completo y concreto (como los paquetes de Ubuntu, mostrado en la Figura 1, o Dockerhub, que ofrece contenedores para Docker en cualquier sistema compatible). En el lado del cliente (el usuario en su ordenador), un administrador de paquetes ayuda a buscar, instalar y actualizar el software disponible desde los repositorios

(ver Figura 1, por ejemplo). En el lado del servidor (el ordenador donde se almacena el software listo para descargarse, es decir, el repositorio) los administradores también utilizan un software específico de repositorios para gestionar los programas ofrecidos a los usuarios: por ejemplo, pueden añadir una nueva versión de un programa para que los usuarios puedan actualizarlo desde casa.

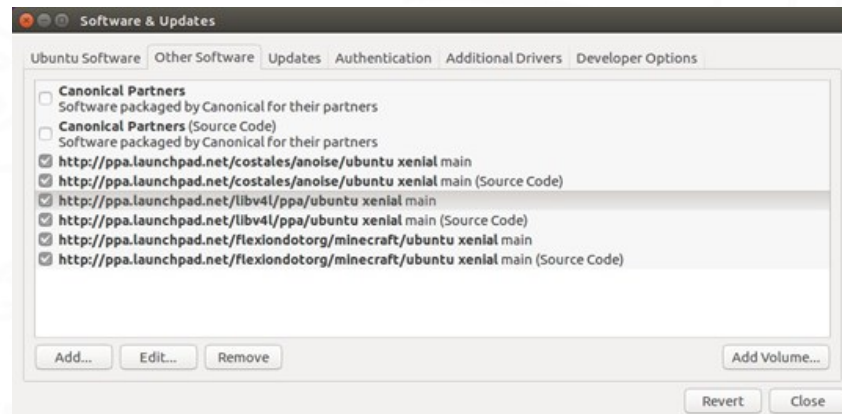


Figura 1: Añadiendo repositorios a Ubuntu. Los paquetes (software) se descargan de ellos.

Fuente: <https://ubuntu.dokry.com/16090/cuales-son-los-repositorys-predeterminados-en-ubuntu-16-04-lts-de-64-bits.html>

### 3.1.4. CÓDIGO FUENTE

Un repositorio de código fuente almacena las distintas versiones de código de una aplicación informática o código fuente, generalmente usando un software de control de versiones. Esto permite a los desarrolladores de software trabajar a la vez en un mismo código de manera coordinada. Sin embargo, también pueden usarse para trabajar en cualquier tipo de fichero, no solo código fuente.

Generalmente las empresas tienen un repositorio disponible donde el software que administra el repositorio permite acceder a él de forma remota y sincronizada para descargar las últimas versiones del código fuente, y subir los cambios realizados por las personas que trabajan en él.

Muchos de estos repositorios son privados para los miembros de una empresa, pero también hay sitios web que ofrecen el servicio de forma gratuita o de pago,

pero permiten que cualquiera pueda acceder al código fuente. *Git* es un ejemplo de software de control de versiones, que permite crear repositorios de código fuente tanto de forma local (en un único ordenador), como de forma remota. *Github* es una web que usa *Git* para ofrecer repositorios a las personas que estén interesadas. Estos repositorios pueden ser públicos (cualquiera puede acceder al código fuente que almacenes en ese repositorio) o privados (sólo se pueden consultar por contraseña)<sup>1</sup>. Otros ejemplos de software de control de versiones para crear repositorios de código son *Subversion*, *CVS* o *Mercurial*, y algunos sitios web que utilizan software de control versiones y permiten usarse de forma gratuita, aparte de *Github*, son *Bitbucket* (<https://bitbucket.org/>) *SourceForge* (<https://sourceforge.net/>).

### 3.1.5. OTROS TIPOS DE DATOS

Existen repositorios que ofrecen otros tipos de datos, como filmes, videos o imágenes. Un ejemplo es *CaCoCu* (Canal de Cultura Contemporánea) de las Universidades andaluzas (<https://www.cacocu.es/>), que ofrece materiales completamente descritos y que pueden utilizarse tanto a nivel personal como educativo, poniendo así a disposición de los visitantes una importante fuente de recursos educativos y sirviendo como un foco de difusión de las distintas expresiones culturales promovidas por las universidades públicas de Andalucía.

Otro ejemplo es *Wikipedia Commons* (<https://commons.wikimedia.org/>), que es un repositorio de archivos multimedia que pone a disposición de todo el mundo, contenidos educativos de dominio público y de licencia libre (imágenes, sonido y videoclips). Otro ejemplo de repositorio de imágenes y música es *Pixabay* (<https://pixabay.com/es/>), que ofrece más de 1.8 millones de archivos de imagen y audio con una licencia libre propia (*Pixabay License*, aunque ha usado *Creative Commons* hasta 2019).

---

<sup>1</sup> En la sección 5 se amplía la información sobre *Git* y *Github*

## 3.2 POR LA TEMÁTICA DE LOS DATOS ALMACENADOS

En este tipo de clasificación entrarían los llamados repositorios temáticos. En lugar de clasificar por tipo de documento, como en la anterior clasificación, los repositorios temáticos se diferencian entre sí, por la temática de los datos almacenados. Por ejemplo, un repositorio de recursos didácticos de matemáticas (que contienen pdfs, imágenes, exámenes resueltos, etc.). Concretamente, este tipo de repositorios se denominan repositorios de objetos de aprendizaje [2].

Algunos de los repositorios temáticos más importantes son:

- Arxiv: (<http://www.arxiv.org>), que almacena artículos científicos de varias disciplinas.
- Eprints (<https://www.eprints.org/>).
- Biogrid, mencionado anteriormente, está limitado a datos de interacciones genéticas.
- ClinicalTrials (<http://www.clinicaltrials.gov>), por ejemplo, ofrece estudios clínicos.

Como ejemplos de repositorios de objetos de aprendizaje tenemos MERLOT (<https://www.merlot.org/merlot/>) o Connexions (<https://cnx.org/>), ambos con licencias abiertas.

## 3.3 POR LA FUENTE DE LOS DATOS GENERADOS

En esta clasificación los repositorios se definen a partir de la fuente de los documentos, es decir, quien es el generador de los documentos que se ofrecen en el repositorio. Dentro de esta clasificación podemos encontrar tres tipos, los institucionales, los empresariales y los generales.

### 3.3.1 INSTITUCIONALES

Los repositorios institucionales son los que han sido creados por las propias instituciones (como las Universidades) para difundir y poner en valor la información que ellas mismas generan. Por ejemplo, el repositorio de producción científica de cualquier universidad, para dar a conocer sus investigadores y su trabajo, o el repositorio institucional de Salud de Andalucía.

Otro ejemplo es la Comisión Europea, que ofrece distintos repositorios. Por

ejemplo, el Repositorio de Discursos (*Speech Repository*) (<https://webgate.ec.europa.eu/sr/>), que es un banco de discursos en el que se pueden encontrar cientos de videos especialmente seleccionados de discursos de la vida real y material pedagógico hecho a medida. El Repositorio de Discursos es un recurso gratuito exclusivamente para fines no comerciales.

La mayoría de estos repositorios están soportados por Software Libre, es decir, el software utilizado para mostrar por internet los datos, y filtrar y buscar en ellos es libre [5] y muchos de los contenidos tienen licencias libres donde se han cedido algunos derechos para poder usarlos.

### 3.3.2 EMPRESARIALES

Este tipo de repositorio están mantenidos por empresas y sirven para difundir su producción. Normalmente estos repositorios suelen estar cerrados al público general, solamente teniendo acceso los miembros de la empresa. Por ejemplo, los repositorios de código o los de paquetes software de una empresa de programación que usa software privativo. Sin embargo, algunas empresas ofrecen estos repositorios también al público en general, por ejemplo, aquellas empresas que trabajan con software Libre, como Mozilla, que ofrece su código fuente en Github (<https://github.com/mozilla>) o Apache, que ofrece todas las versiones de su software listo para descargar.

### 3.3.3 GENERALES

Estos repositorios recopilan información de muchas fuentes distintas, no necesariamente producidas por la propia entidad que mantiene el repositorio. Por ejemplo, Figshare (que explicaremos más adelante) está mantenido por la empresa Digital Science, permitiendo que cualquier persona del mundo suba sus datos, imágenes artículos, etc. Otros ejemplos similares son Dryad (<https://datadryad.org/stash>) de la Universidad de Oxford o Zenodo (<https://zenodo.org/>), gestionado por el CERN (Organización Europea para la investigación Nuclear). Como puede verse, estos repositorios están sobre todo orientados al ámbito científico.



## 3.4 POR EL TIPO DE APERTURA DE ACCESO

La mayoría de repositorios disponibles en internet tienen como misión compartir sus contenidos. Sin embargo, puede ser necesario también mantener un histórico de documentos y datos dentro de una misma organización. Un ejemplo de esto son los repositorios documentales dentro de los departamentos de cada Universidad, donde se almacena información sensible que puede ser consultada solamente por los miembros del departamento, como las actas de una reunión. Es por ello que podemos definir dos tipos de modalidades de acceso.

### 3.4.1 REPOSITARIOS ABIERTOS

Son aquellos que permiten acceder a los datos a cualquier persona a través de internet, normalmente ofreciendo alguna licencia abierta a los datos ofrecidos, como, por ejemplo, Creative Commons.

En el Registry of Open Access Repositories pueden encontrarse muchos repositorios Open Access (<http://roar.eprints.org/>), permitiendo realizar una búsqueda por tipo, fecha, país, etc.

### 3.4.2 REPOSITARIOS CERRADOS

Son aquellos cuyos datos privados únicamente accesibles para aquellos miembros de una comunidad o empresa concreta. Normalmente se pueden acceder por internet, pero usando credenciales (identificador y contraseña)

## 4. EJEMPLOS DE REPOSITARIOS

En esta sección describiremos algunos repositorios existentes ampliamente utilizados

### 4.1 ARXIV

arXiv es un repositorio de acceso abierto de preimpresiones electrónicas (conocidas como *e-prints*) aprobadas para su publicación con moderación, pero no para su revisión completa por pares [1]. Comprende documentos científicos en los campos de las matemáticas, la física, la astronomía, la

ingeniería eléctrica, la informática, la biología cuantitativa, la estadística, las finanzas matemáticas y la economía, a los que se puede acceder en línea. En muchos campos de las matemáticas y la física, casi todos los artículos científicos se archivan en el repositorio de arXiv antes de su publicación en una revista revisada por pares. Iniciado el 14 de agosto de 1991, arXiv.org superó el hito del medio millón de artículos el 3 de octubre de 2008, y ha alcanzado el millón a finales de 2014. Para octubre de 2016 la tasa de envío había crecido a más de 10.000 por mes. Permite hacer búsquedas en su página web (ver Figura 2), y como otros muchos repositorios ofrece un acceso a una API de consulta para obtener los metadatos de los documentos (Figura 3).

## 4.2 DIALNET

Dialnet es un portal de difusión de la producción científica hispana cuyo funcionamiento se inició en 2001 especializado en ciencias humanas y sociales. El portal está gestionado por la Fundación Dialnet, de la Universidad de La Rioja, una entidad sin ánimo de lucro creada en febrero de 2009 para la gestión y desarrollo de una de las mayores bases de datos de literatura científica del mundo. Incluye libros (monografías), tesis doctorales, congresos, homenajes y otro tipo de documentos. El texto completo de muchos de los documentos está disponible en línea. En el portal colaboran bibliotecas de numerosas universidades españolas e hispanoamericanas y bibliotecas públicas y especializadas que realizan los volcados de sumarios de revistas.

También incorpora base de datos con documentos en otros idiomas [7].

En el ranking del Laboratorio de Cibermetría del CSIC, Dialnet ocupa el primer puesto entre los portales europeos y el cuarto a nivel mundial.

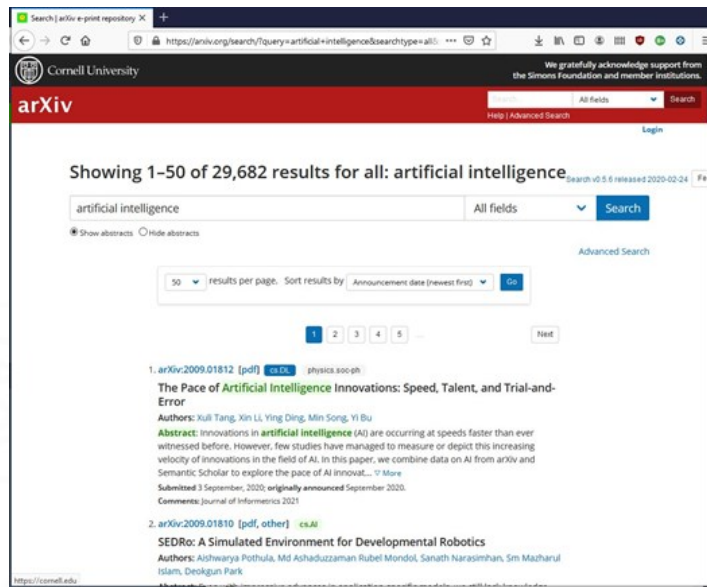


Figura 2: Ejemplo de búsqueda en archivo Arxiv. Como puede verse, se puede descargar el fichero PDF del artículo buscado.

### 4.3 FIGSHARE

Figshare (<https://figshare.com/>) es un repositorio de acceso abierto en línea en el que los investigadores pueden preservar y compartir sus resultados de investigación, incluyendo documentos, figuras, conjuntos de datos, imágenes y vídeos. Es de libre acceso y de subida gratuita, en adhesión al principio de datos abiertos. Está mantenida por la empresa Digital Science.

Los usuarios pueden subir archivos en cualquier formato, y a los artículos se les atribuye un DOI. Todos los archivos son liberados bajo una licencia Creative Commons, CC-BY para la mayoría de los archivos y CC0 (dominio público) para los conjuntos de datos. Puede verse una captura de su web en la Figura 5. También permite la publicación de datos negativos (que han salido mal). Al fomentar la publicación de figuras, gráficos y datos, en lugar de limitarse al tradicional “artículo” completo, el conocimiento puede ser compartido más rápida y eficazmente. Figshare también hace un seguimiento de las estadísticas de descarga de los materiales alojados, actuando a su vez como fuente de *altmetrics* (o métricas alternativas al tradicional número de

citas de un artículo, como por ejemplo el número de descargas o menciones en twitter). Los datos se almacenan en Amazon S3 y CLOCKS, permitiendo que los datos estén respaldados y distribuidos [8].

```

<feed>
<link href="http://arxiv.org/api/query?search_query%3Dall%3Aelectron%26id_list%3D%26start%3D0%26max_res
<title type="html">
ArXiv Query: search_query=all:electron&id_list=&start=0&max_results=10
</title>
<id>http://arxiv.org/api/WyBPOs+pRgzCTXTMWhmbcOmk6g</id>
<updated>2020-09-04T00:00:00-04:00</updated>
<opensearch:totalResults>161864</opensearch:totalResults>
<opensearch:startIndex>0</opensearch:startIndex>
<opensearch:itemsPerPage>10</opensearch:itemsPerPage>
<entry>
<id>http://arxiv.org/abs/cond-mat/0102536v1</id>
<updated>2001-02-28T20:12:09Z</updated>
<published>2001-02-28T20:12:09Z</published>
<title>
Impact of Electron-Electron Cusp on Configuration Interaction Energies
</title>
<summary>
The effect of the electron-electron cusp on the convergence of configuration interaction (CI) wave functions is e
scattering of the Coulomb interaction but is smooth and finite at zero electron-electron separation. The exact ma
and Be atoms, both with the Coulomb electron-electron interaction and with the smooth effective electron-electri
that for the divergent Coulomb interaction for energy differences on the order of 1 mHartree. This shows that, co
</summary>
<author>
<name>David Prendergast</name>
<arxiv:affiliation>Department of Physics</arxiv:affiliation>
</author>
<author>
<name>M. Nolan</name>
</author>

```

Figura 3: Ejemplo de metadatos de Arxiv.

Estos datos pueden obtenerse automáticamente desde un programa externo en cualquier lenguaje de programación

## 5. SOFTWARE LIBRE PARA REPOSITORIOS

En esta sección definiremos brevemente algunos ejemplos de software libre para montar un repositorio. En [5] puede verse una comparativa de algunos de estos ejemplos.

### 5.1 GIT

Como se ha dicho anteriormente, uno de los sistemas de control de versiones más conocidos es Git. Git es un sistema distribuido de control de versiones

para rastrear los cambios en el código fuente durante el desarrollo del software. Sus objetivos incluyen la velocidad, la integridad de los datos, y el apoyo a los flujos de trabajo distribuidos y no lineales [9].

Git no debe confundirse con Github. Git es el software para repositorios de código, mientras que Github es una plataforma web, actualmente propiedad de Microsoft, que utiliza Git para ofrecer repositorios a las personas que los necesiten y no quieran montar toda la infraestructura que se necesita para gestionar un repositorio global.

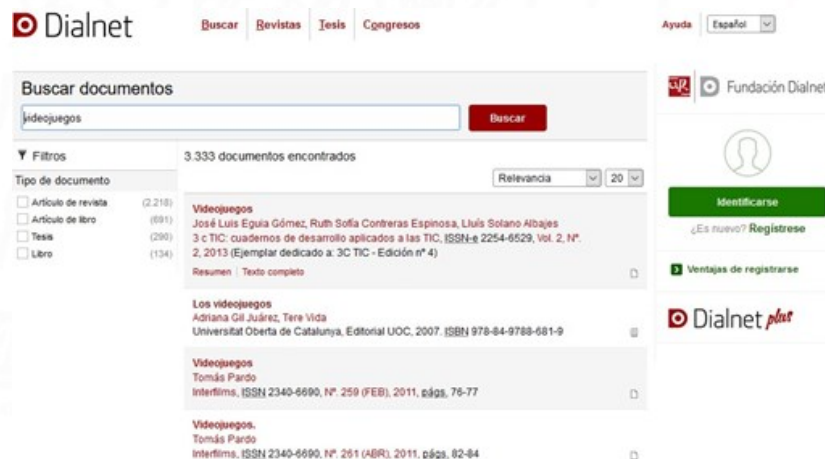


Figura 4: Búsqueda en DialNet

## 5.2 ALFRESCO

Alfresco es una colección de programas de gestión de la información desarrollados por Alfresco Software Inc. utilizando Java. Su principal oferta de software es una plataforma de código abierto propietaria y con licencia comercial, compatible con estándares abiertos. Este repositorio sirve para la gestión de contenidos empresariales para documentos, web, registros, imágenes, vídeos, medios enriquecidos y desarrollo de contenidos de forma colaborativa (ver Figura 7). Normalmente se usa dentro de empresas, de forma cerrada, solamente permitiendo el acceso a los miembros de la empresa.

Sin embargo, Alfresco Software Inc. también proporciona la edición Community de este repositorio, con licencia LGPLv3. Tiene algunas restricciones predeterminadas en términos de escalabilidad y disponibilidad, pero permiten a cualquiera la instalación y modificación del código. La versión Community también está muy extendida a nivel empresarial.

### 5.3 EPRINTS

Eprints, mencionado anteriormente, también ofrece el software que han desarrollado para que se instale de forma independiente en cualquier ordenador. Tiene una licencia GNU. Eprints se puede instalar desde los repositorios de Ubuntu y Fedora, pero también se puede obtener también desde su página de Github (<https://github.com/eprints/eprints3.4>).

### 5.4 DSPACE

DSpace (disponible en <https://github.com/DSpace/DSpace/> con una licencia BSD) también puede usarse como solución de repositorio bibliográfico institucional. Cada ítem pertenece a una colección, y cada colección a una comunidad. Al igual que otro software, su cometido está relacionado con administración de registros electrónicos, la preservación digital y la publicación.

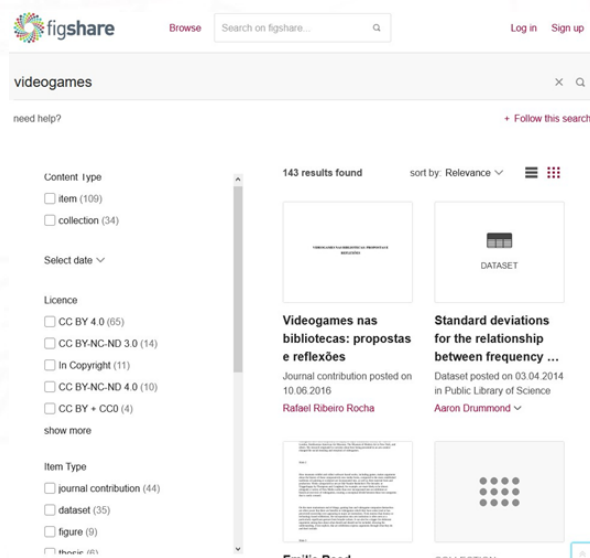


Figura 5: Búsqueda de Figshare. Puede verse cómo se puede filtrar por tipo, licencia, etc

## 5.5 INVENIO

Invenio (antes conocido como CDS) es otro software de código abierto para repositorios digitales en gran escala que proporciona las herramientas para la gestión de activos digitales en un repositorio institucional y sistemas de gestión de datos de investigación. Fue desarrollado inicialmente por el CERN. El código, está disponible en <https://github.com/inveniosoftware>.

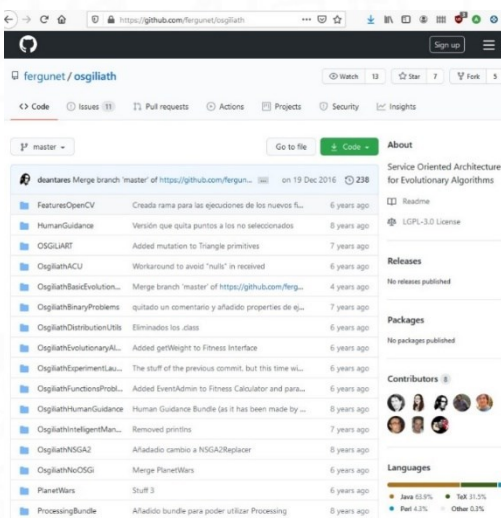


Figura 6: Ejemplo de un repositorio Git en Github

## REFERENCIAS:

- [1] Paul Ginsparg. Arxiv at 20. *Nature*, 476(7359):145–147, 2011.
- [2] R.E. Jones, T. Andrew, and J. MacColl. *The Institutional Repository*. Chandos Information Professional Series. Elsevier Science, 2006.
- [3] Chris Stark, Bobby-Joe Breitreutz, Teresa Reguly, Lorrie Boucher, Ashton Breitreutz, and Mike Tyers. Biogrid: a general repository for interaction datasets. *Nucleic acids research*, 34(suppl1):D535–D539, 2006.
- [4] Daniel Torres-Salinas, Nicolás Robinson-García, and Álvaro Cabezas-Clavijo. Compartir los datos de investigación en ciencia: introducción al data sharing. *El profesional de la información*, 21(2):173-184, 2012
- [5] Jesús Tramullas and Piedad Garrido. Software libre para repositorios institucionales: Propuestas para un modelo de evaluación de prestaciones. *El Profesional de la Información*, 15(3):171–181, 2006.
- [6] Herbert Vande Sompel, Michael Nelson, Carl Lagoze, and Simeon Warner. Resource harvesting within the oai-pmh framework. *D-lib magazine*, 10(12), 2004.
- [7] Wikipedia. Dialnet. <https://en.wikipedia.org/wiki/Dialnet>, 2020. [Online; accessed 4-September-2020].
- [8] Wikipedia. Figshare. <https://en.wikipedia.org/wiki/Figshare>, 2020. [Online; accessed 4-September-2020].
- [9] Wikipedia. Git. <https://en.wikipedia.org/wiki/Git>, 2020. [Online; accessed 4-September-2020].



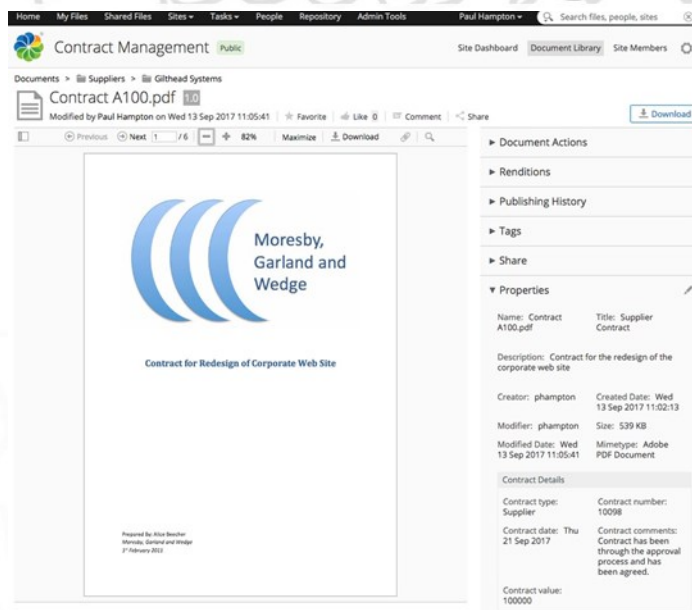


Figura 7: Ejemplo de un documento disponible en un repositorio Alfresco. Puede verse el documento, junto con sus metadatos. Imagen obtenida de: [https://en.wikipedia.org/wiki/Alfresco\\_Software](https://en.wikipedia.org/wiki/Alfresco_Software).

[10] Wikipedia. JSON. <https://es.wikipedia.org/wiki/JSON>, 2020. [Online; accessed 4-September-2020].

[11] Wikipedia. Software Repositories. [https://en.wikipedia.org/wiki/Software\\_repository](https://en.wikipedia.org/wiki/Software_repository), 2020. [Online; accessed 4-September-2020].