

GUÍA PARA LA EVALUACIÓN DE REPOSITORIOS INSTITUCIONALES DE INVESTIGACIÓN



GOBIERNO
DE ESPAÑA

MINISTERIO
DE ECONOMÍA
Y COMPETITIVIDAD



FECYT

FUNDACIÓN ESPAÑOLA
PARA LA CIENCIA
Y LA TECNOLOGÍA

Recolecta

RECOLECTOR DE CIENCIA ABIERTA



CRUE

COORDINADORES

Barrueco Cruz, José Manuel

Universitat de València

De Miguel Estévez, Mercedes

Fundación Española para la Ciencia y la Tecnología

González Copeiro, Cristina

Fundación Española para la Ciencia y la Tecnología

Rico-Castro, Pilar

Fundación Española para la Ciencia y la Tecnología

AUTORES

Azorín Millaruelo, Cristina

Universitat Autònoma de Barcelona

Cívico Martín, Rafaela

Universidad de Huelva

Combarro Felpeto, Pilar

Universidade da Coruña

García Gil, M^a Ángeles

Universidad de Granada

Losada Yáñez, Marina

Universitat Pompeu Fabra

Moreno Cañizares, Antonio

Universidad Complutense de Madrid

Morillo Moreno, José Carlos

Universidad de Huelva

Pérez Iglesias, Javier

Universidad Complutense de Madrid

Serrano Vicente, Rocío

Universidad de Navarra

Esta es una modificación de la segunda versión de junio de 2014 de la *Guía para la evaluación de repositorios institucionales de investigación*, cuya primera edición fue publicada por RECOLECTA en Diciembre de 2010.



SUMARIO

1.- INTRODUCCIÓN	4
2.- DIRECTRICES TENIDAS EN CUENTA	7
3. - FINALIDAD DE LA GUÍA. A QUIÉN VA DIRIGIDA	8
4.- ESTRUCTURA	10
5. CRITERIOS RECOLECTA PARA LA EVALUACIÓN DE REPOSITARIOS INSTITUCIONALES DE INVESTIGACIÓN	11
1.- VISIBILIDAD	11
2.- POLÍTICAS	13
3.- ASPECTOS LEGALES	14
4.- METADATOS	15
5.- INTEROPERABILIDAD	19
6.- LOGS Y ESTADÍSTICAS	21
7.- SEGURIDAD, AUTENTICIDAD E INTEGRIDAD DE LOS DATOS	21
6.- GLOSARIO	22
ANEXO 1 Vocabulario establecido para el campo tipo de publicación (dc:type)	26
ANEXO 2 Vocabulario establecido para indicar la versión del documento en el campo tipo de publicación (dc:type)	28
ANEXO 3 Vocabulario establecido para el campo tipo de formato (dc:format)	29
ANEXO 4 Vocabulario establecido para el campo nivel de acceso (dc:rights)	30
ANEXO 5 Criterios RECOLECTA para contenido	31
ANEXO 6 Criterios RECOLECTA obligatorios para protocolo OAI-PMH	32
BIBLIOGRAFÍA	33
WEBGRAFÍA	34



1.- INTRODUCCIÓN

El **acceso abierto** (*Open Access*) consiste en proveer de acceso on-line a toda la información científica disponible (artículos, monografías, datos de investigación,...) de forma gratuita para el lector y bajo licencia para que pueda ser usada y explotada por los investigadores, las empresas y los ciudadanos¹. La idea que subyace bajo el acceso abierto es que la investigación financiada públicamente debería estar disponible para todos.

La investigación se hace más efectiva y sus resultados más visibles cuando se eliminan las barreras legales, comerciales y tecnológicas de acceso a la información científica. Además, el acceso abierto previene la duplicación, fomenta la transferencia tecnológica y de conocimiento y promueve la innovación. Todos los actores implicados en el sistema de comunicación de la información científica se benefician del acceso abierto a la investigación y a los datos generados por ésta.

El reciente interés que se ha suscitado mundialmente por las políticas de acceso abierto es resultado de muchos años de trabajo promocionando este nuevo modelo de acceso a la información científica por parte de los investigadores y los defensores de esta política.

El avance más relevante que se está produciendo en este contexto es el creciente número de entidades financiadoras, universidades y centros de investigación que actualmente cuentan con **políticas** de acceso abierto en el mundo. Las agencias de financiación de la investigación científica más relevantes, públicas y privadas, están implantando políticas de acceso abierto obligatorias, construyendo así los cimientos para que el acceso abierto sea el modo habitual de comunicar la investigación. Igualmente, importantes universidades y centros de investigación están dando los pasos necesarios en esta dirección.

En Europa, el programa Horizonte 2020 incluye entre las obligaciones de sus beneficiarios la de depositar en abierto todas aquellas publicaciones que se generen en el marco de los proyectos financiados. Además, la Comisión Europea ha lanzado un programa piloto que afecta a 7 áreas de conocimiento que hace obligatorio el depósito en abierto de los datos de investigación producidos gracias a estos proyectos². En España, la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la

¹ La definición de acceso abierto se recoge en la *Iniciativa de Budapest para el Acceso Abierto* (<http://www.budapestopenaccessinitiative.org/read>), de 2002, y en la *Declaración de Berlín sobre el Acceso Abierto* (<http://openaccess.mpg.de/2365/en>), de 2003.

² Las áreas incluidas en el piloto de datos son: 1. Future and Emerging Technologies; 2. Research infrastructures – part e-Infrastructures; 3. Leadership in enabling and industrial technologies – Information and Communication Technologies; 4. Societal Challenge: Secure, Clean and Efficient Energy – part Smart cities and communities; 5. Societal Challenge: Climate Action, Environment, Resource Efficiency and Raw materials – with the exception of topics in the area of raw materials; 6. Societal Challenge: Europe in a changing world – inclusive, innovative and reflective Societies; 7. Science with and for Society. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf



Innovación, en su artículo 37, obliga a los beneficiarios de proyectos de investigación financiados por Plan Estatal de Investigación Científica y Técnica y de Innovación a depositar en repositorios de acceso abierto la versión digital final de las contribuciones a publicaciones periódicas o seriadas. Además, hay tres gobiernos regionales que han tomado la iniciativa y han desarrollado sus propias regulaciones sobre el depósito en abierto: la Comunidad de Madrid, el Principado de Asturias y la Generalitat Catalana. Por último, el depósito en abierto también ha recibido un importante impulso a nivel institucional y actualmente en España hay 26 instituciones de investigación que han desarrollado su propia política institucional en favor del acceso abierto, ya sea en forma de declaración institucional, recomendación o de requisito de obligado cumplimiento³.

El surgimiento de estas políticas requiere contar con las **infraestructuras** necesarias para que los investigadores puedan dar cumplimiento a la obligación de depositar en abierto. Así, de forma paralela al desarrollo de mandatos y al diseño e implementación de políticas de acceso abierto, se están produciendo importantes avances en materia de infraestructuras gracias al desarrollo de las tecnologías de la información y la comunicación. Los repositorios y las revistas de acceso abierto constituyen los recursos electrónicos necesarios para poder ejecutar los mandatos de acceso abierto.

Las revistas científicas de acceso abierto son aquellas que permiten el acceso inmediato a sus contenidos, sin restricciones de tipo técnico o económico. Normalmente los autores conservan sus derechos de autor. Las publicaciones en acceso abierto siguen los mismos procesos que las publicaciones comerciales, tales como la evaluación por pares (también conocida como *peer review*), pero proveen de acceso abierto directo a sus contenidos.

Respecto a los repositorios de acceso abierto, hay tres tipos distintos: Los repositorios institucionales son infraestructuras ubicadas en las universidades y centros de investigación, normalmente en las bibliotecas, que permiten a la institución gestionar, preservar y mostrar su producción científica. Los repositorios temáticos son aquellos que recogen la producción de determinadas áreas de conocimiento a nivel internacional (ejemplos: Arxiv.org, RePEc, PsyDok, PubMed). Los repositorios centralizados son aquellos que recogen la producción científica depositada en abierto por investigadores de distintas instituciones y de distintas áreas temáticas (ejemplo: Zenodo, Reposit-AN, Recercat, TDR).

En España existe una importante infraestructura de recursos en los que depositar en abierto los trabajos de investigación, dentro de la que destaca la comunidad de repositorios institucionales que trabajan para hacer posible el acceso libre y gratuito a toda la producción científica de sus

³ Consejo Superior de Investigaciones Científicas, Universidad de Alcalá, Universidad de Alicante, Universitat Autònoma de Barcelona, Universitat de Barcelona, Universidad Carlos III de Madrid, Universidad de Cantabria, Universidad CEU Cardenal Herrera, Universidad Complutense de Madrid, Universidad de Extremadura, Universitat de Girona, Universidad de León, Universitat de Lleida, Universidad de Málaga, Universitat Oberta de Catalunya, Universidad Politécnica de Cartagena, Universitat Politècnica de Catalunya, Universidad Politécnica de Madrid, Universitat Politècnica de València, Universitat Pompeu Fabra, Universidad Rey Juan Carlos, Universidad de Salamanca (tesis), Universidade de Santiago de Compostela, Universidad de Valladolid (tesis), Universitat de Vic, Universidad de Zaragoza. FUENTE: <http://recolecta.fcyc.es> y www.accesoabierto.es [Fecha de consulta: 30 de mayo de 2014].



universidades y centros de investigación. Hoy en día aproximadamente el 80% de las universidades españolas disponen de su propio repositorio institucional.

La interoperabilidad de todos estos repositorios y la unificación de criterios de calidad resulta ser de gran importancia para lograr el máximo aprovechamiento de la información que contienen. Por ello, desde 2007 la **Fundación Española para la Ciencia y la Tecnología (FECYT)** y la **Red de Bibliotecas Universitarias (REBIUN)** de la CRUE colaboran en el proyecto **RECOLECTA** (o *Recolector de Ciencia Abierta*). Los principales objetivos de RECOLECTA son los siguientes:

- Impulsar y coordinar la infraestructura nacional de repositorios científicos digitales de acceso abierto y garantizar que sean interoperables según los estándares de la comunidad mundial.
- Promover, apoyar y facilitar la adopción del acceso abierto por todos los investigadores de las universidades y centros de I+D españoles, principales productores de conocimiento científico en nuestro país.
- Dotar de una mayor visibilidad tanto nacional como internacional de los resultados de la investigación que se realiza en España.

Las revistas y repositorios de acceso abierto españoles están agrupados dentro de la plataforma de RECOLECTA (<http://recolecta.fecyt.es>) de manera que, a través de ella, se puede acceder de forma centralizada a toda la producción publicada en abierto, tanto en revistas como en repositorios institucionales, temáticos y centralizados. Además, RECOLECTA garantiza la interoperabilidad de todos los repositorios y trabaja para proveer a la comunidad de servicios de valor añadido, tales como estadísticas de visitas y descargas.

Para alcanzar un elevado estándar de calidad, la presente versión de la *Guía para la evaluación de repositorios institucionales de investigación* ofrece a toda la comunidad nacional un conjunto de directrices, basadas en criterios internacionales ya existentes, que garantizan la interoperabilidad de todos los recursos de acceso abierto y aseguran un acceso de calidad a sus contenidos. Además, en esta guía también se recogen los criterios obligatorios que marca RECOLECTA para ingresar en su plataforma⁴.

⁴ En el siguiente enlace se pueden consultar los criterios RECOLECTA:
http://recolecta.fecyt.es/sites/default/files/contenido/documentos/CRITERIOS_RECOLECTA_DRIVER.pdf



2.- DIRECTRICES TENIDAS EN CUENTA

Para la elaboración de la presente guía se ha realizado un exhaustivo análisis de las iniciativas que han tratado de establecer criterios de calidad comunes a toda la comunidad internacional de repositorios de acceso abierto. Por ello, esta guía contiene una panorámica actualizada de estas directrices, aderezada con la experiencia de la comunidad RECOLECTA en gestión de repositorios institucionales.

Se ha utilizado como principal referencia OpenAIRE (*Open Access Infrastructure for Research in Europe*) que ha absorbido al antiguo DRIVER (*Digital Repository Infrastructure Vision for European Research*). Como consecuencia de esta absorción, se han fusionado sus infraestructuras. Además, las directrices OpenAIRE se han construido sobre las antiguas directrices DRIVER. Éstas se refieren a los siguientes aspectos:

- Implementación de OAI-PMH: define características obligatorias y recomendadas para solucionar problemas en distintas implementaciones de repositorios de acceso abierto.
- Recursos textuales: contempla los requerimientos que deben cumplir los recursos. Destacamos la obligación de utilizar “sets” (agrupaciones) que definan las colecciones accesibles a texto completo.
- Metadatos: marca y define los elementos Dublin Core obligatorios y recomendados.

Durante los últimos años, OpenAIRE ha elaborado diversos paquetes de recomendaciones: *Guidelines OpenAIRE 1.1*, *Guidelines OpenAIRE 2.0*, *Guidelines OpenAIRE 2.2* y *Guidelines OpenAIRE 3.0*. Como norma general, las directrices de OpenAIRE han mantenido el corpus de directrices DRIVER. Sin embargo, hay excepciones relevantes que han sido recogidas en esta guía (por ejemplo, la información acerca de los derechos contenidos en y sobre el recurso (dc:rights) era recomendado en DRIVER y es obligatorio en OpenAIRE).

Además de las directrices mencionadas, esta guía ofrece pautas concretas para ahondar en diversos aspectos políticos, legales y técnicos que son de gran importancia para la comunidad de repositorios que albergan contenidos científicos digitales y que quieren acceder a servicios globales de búsqueda y localización de contenidos.

Respetar estos estándares de compatibilidad e interoperabilidad hace posible que el contenido de un repositorio sea más fácilmente recuperable y visible, no sólo para la propia institución sino para la comunidad científica a nivel global y la sociedad en general.



3.- FINALIDAD DE LA GUÍA. A QUIÉN VA DIRIGIDA

Esta segunda versión de la *Guía para la evaluación de repositorios institucionales de investigación* tiene como finalidad principal ser un instrumento de auditoría interna para mejorar la calidad de los repositorios. Disponer de esta herramienta de autoevaluación supone una importante ayuda a la hora de detectar puntos fuertes y áreas de mejora, lo cual contribuye a incrementar el alcance y la visibilidad de los repositorios entre la comunidad científica. Además, para los repositorios que se encuentran en fase de planificación, este documento ofrece las pautas necesarias a tener en cuenta para crear una infraestructura robusta y visible.

Esta guía está dirigida primordialmente a los profesionales implicados en la gestión de repositorios científicos: responsables, administradores, gestores, técnicos o personal de apoyo. Para elaborarla se ha constituido un grupo de trabajo, en el marco de la colaboración FECYT-REBIUN, con representantes de las universidades que participaron en la primera edición del documento en 2010.⁵

Este grupo de trabajo define como repositorio institucional de investigación “aquél conjunto de servicios prestados por las universidades o centros de investigación a su comunidad para recopilar, gestionar, difundir y preservar su producción científica digital a través de una colección organizada, de acceso abierto e interoperable”.

Un repositorio puede tener una estructura muy variada y ubicar sus colecciones de forma muy diversa. Dado que un repositorio contiene, o puede contener, todos los documentos de la institución que lo alberga, la variabilidad de dichos registros es muy amplia. Así, por ejemplo, el repositorio puede contener materiales docentes, vídeos, fondo antiguo, colecciones patrimoniales, colección de investigación, etc.

A su vez, dentro de cada uno de esos grupos existe la opción de crear *sets* o grupos homogéneos en los que se agrupen determinados elementos de una colección o en el que se relacionen elementos de diferentes colecciones. Esta posibilidad resulta de extrema utilidad a la hora de recolectar sólo determinados *sets* de un repositorio. **Esta guía está dirigida únicamente a la colección de investigación del repositorio.**

En el siguiente esquema se expone gráficamente la estructura general de un repositorio:

⁵ El grupo de trabajo está formado por las siguientes universidades: Universitat Autònoma de Barcelona, Universidade da Coruña, Universidad de Huelva, Universidad de Granada, Universidad Complutense de Madrid, Universitat Pompeu Fabra, Universidad de Navarra y Universitat de València.

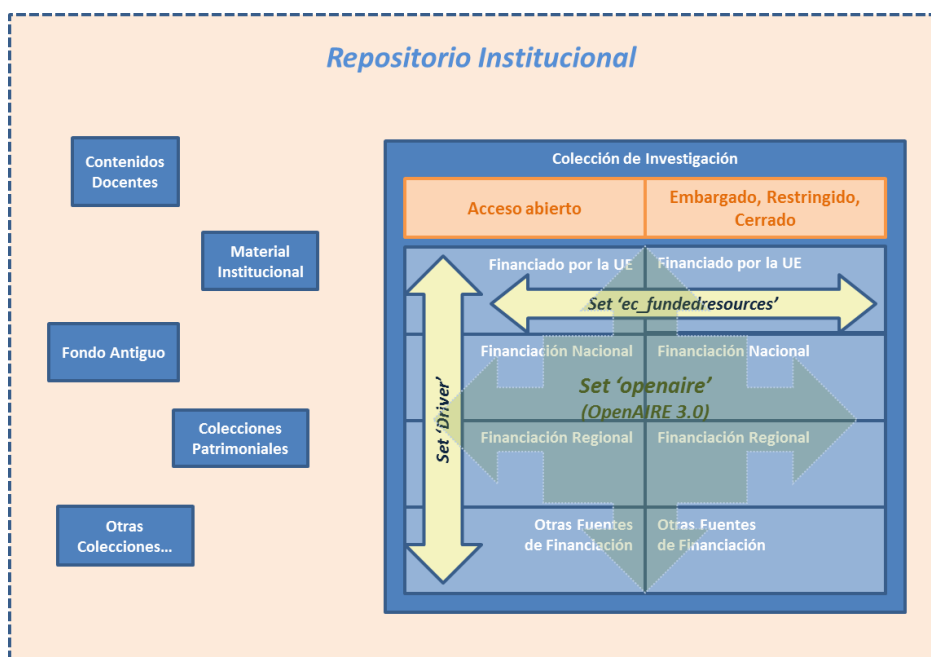


Gráfico 1: Estructura general de un repositorio institucional. Fuente: Elaboración propia

Los repositorios objeto de esta guía deben cumplir con unos requisitos mínimos que se relacionan a continuación:

- **Recopilar, gestionar, difundir y preservar la producción científica digital de la institución.** Los documentos almacenados en el repositorio deben haber sido creados por la institución o alguno de sus miembros como producto de las funciones de investigación que le son propias. Si en el repositorio se mezclan diferentes tipos de colecciones (científica, docente, cultural, institucional), cada una de ellas deberá estar claramente diferenciada tanto en el repositorio como en los metadatos que se pegan a disposición de recolectores.
- **Colección organizada.** Un repositorio no puede ser un mero depósito de documentos. Estos deberán estar descritos utilizando un número suficiente de metadatos basados en estándares internacionales y mínimamente normalizados, organizados mediante la aplicación de alguna clasificación de contenidos; deben servir para responder a las necesidades de sus usuarios y prestarles servicios de valor añadido, etc. La colección deberá contener un mínimo de documentos para poder ser considerada un servicio en producción más que un proyecto. El crecimiento del contenido del repositorio debe ser continuo.
- **De acceso abierto.** El objetivo de un repositorio debe ser aumentar la visibilidad e impacto de la investigación que se realiza en una institución a través del libre acceso a dicha producción. Por lo tanto, la situación ideal sería aquella en la que el 100% de los contenidos depositados son de libre acceso.
- **Interoperable.** El repositorio no es un fin en sí mismo sino que su verdadero potencial se descubre cuando sus contenidos se integran en un nivel superior de agregación desde donde se puedan prestar servicios especializados a comunidades concretas. Por lo tanto es fundamental la interoperabilidad del repositorio a través de la utilización de protocolos como el OAI-PMH.



4.- ESTRUCTURA

En esta nueva versión de la guía se han incluido nuevos criterios actualizados para adaptar los repositorios españoles a las directrices europeas e internacionales de interoperabilidad. Algunos de ellos se consideran básicos y de obligado cumplimiento para que los repositorios sean incluidos en el portal RECOLECTA (Anexo 5). En cambio, otros tienen la consideración de buenas prácticas para los repositorios y no se exige, sino que se recomienda, su cumplimiento.

La guía se estructura en un total de siete secciones que definen en conjunto 53 criterios de evaluación.

Cada una de las secciones se dedica al análisis de aspectos fundamentales de los repositorios como son:

1. Visibilidad
2. Políticas
3. Aspectos legales
4. Metadatos
5. Interoperabilidad
6. Logs y estadísticas
7. Seguridad, autenticidad e integridad de los datos

Todos los criterios tienen respuestas cerradas con valores *Sí* o *No* que permiten valorar si el repositorio lo cumple o no.

Las reglas de metadatos establecidas como obligatorias se considera que se deben cumplir por parte de todos los registros y no sólo por un porcentaje de ellos.



5.- CRITERIOS RECOLECTA PARA LA EVALUACIÓN DE REPOSITARIOS INSTITUCIONALES DE INVESTIGACIÓN

1.- VISIBILIDAD

En esta sección se analiza la visibilidad del repositorio a través de su presencia en directorios nacionales e internacionales, así como en la existencia de un nombre normalizado que lo identifique en todos ellos.

1.1.- Mención del repositorio en la página principal de la institución.

Existe un enlace desde la página inicial de la institución al repositorio.

Resultado:

- Sí (especificar)
- No

1.2.- Presencia en directorios nacionales e internacionales.

El repositorio ha sido registrado en al menos tres de los siguientes directorios: ROAR, OpenDOAR, BuscaRepositorios, Hispana y OAI Data Providers. Para conseguir la mayor visibilidad se recomienda el registro en todos ellos.

Resultado:

- Sí (especificar)
- No

1.3.- Presencia en recolectores nacionales e internacionales.

El repositorio está siendo recolectado por al menos tres de los siguientes recolectores: Google Scholar, OpenAire, RECOLECTA y BASE.

Resultado:

- Sí (especificar)
- No

1.4.- Existencia de un nombre normalizado en todos ellos.

El repositorio ha sido registrado en directorios y recolectores siempre con la misma forma del nombre. Se valora que el repositorio tenga un nombre propio que lo identifique unívocamente.

Resultado:

- Sí
- No



1.5.- Existencia de una URL amigable.

Se entiende por URL amigable aquella que está compuesta únicamente por la dirección del servidor web. Se valora que en esta dirección aparezca el nombre del repositorio.

Resultado:

- Sí
- No

1.6.- Existencia de iniciativas para fomentar la visibilidad del repositorio dentro de la propia institución.

Entre éstas, se valora la existencia de: una oferta de sesiones de formación e información sobre la introducción de los documentos en el repositorio; acciones de fomento del acceso abierto mediante organización de eventos, presentaciones, campañas en facultades y departamentos, semanas de acceso abierto, seminarios, pósters y cartelería; utilización de los media; utilización de redes sociales, así como guías y materiales de soporte y asesoramiento disponibles para los autores, etc.

Resultado:

- Sí (especificar cuáles)
- No

1.7.- Al menos el 75% de los recursos textuales de investigación que ofrece el repositorio se encuentran en acceso abierto.

Resultado:

- Sí
- No

1.8.- La Institución se ha adherido a alguna de las declaraciones open access (Budapest, Berlín o Bethesda).

Resultado:

- Sí (especificar cuáles)
- No



2.- POLÍTICAS

2.1.- Existe una declaración sobre la misión y objetivos del repositorio.

Existe un documento de acceso público, fácilmente accesible desde la página principal del repositorio en el que se establecen cuáles son los objetivos, alcance y funciones del mismo.

Resultado:

- Sí (añadir url)
- No

2.2.- Documento de acceso público sobre el archivo en el repositorio, donde se establecen al menos los siguientes puntos: quién puede depositar, qué se puede depositar y en qué formatos.

Existe un documento de acceso público, fácilmente accesible desde la página principal del repositorio, en el que se establece de forma clara qué personas dentro de la institución pueden aportar contenidos, qué tipos de contenidos son aceptados (artículos publicados en revistas, informes, etc.) así como los formatos de los ficheros permitidos (PDF, Word, etc.).

Resultado:

- Sí (añadir url)
- No

2.3.- Documento de acceso público sobre preservación de los contenidos.

Existe un documento de acceso público, fácilmente accesible desde la página principal del repositorio, en el que la institución expresa su compromiso en hacer disponibles los contenidos de forma permanente y tomar las medidas de preservación (tales como migraciones) necesarias para garantizar el acceso a los mismos. Siempre que sea posible crear y conservar formatos de archivo con la finalidad de asegurar su preservación.

Resultado:

- Sí (añadir url)
- No

2.4.- Documento de acceso público sobre reutilización de metadatos.

Los metadatos almacenados en el repositorio pueden ser recolectados por agregadores o proveedores de servicios. Existe un documento de acceso público, fácilmente accesible desde la página principal del repositorio, donde se especifica cómo y en qué medida o con qué limitaciones dichos recolectores pueden utilizar los metadatos recolectados.

Resultado:

- Sí (añadir url)
- No



2.5.- Existe una oferta de contacto y asesoramiento visible.

Se valora la existencia de diferentes medios de contacto (redes sociales, correo electrónico, teléfono, etc.) para realizar asesoramiento telefónico y/o personal a los autores.

Resultado:

- Sí (especificar)
- No

2.6.- Política institucional sobre acceso abierto.

El Repositorio refleja en un lugar visible su compromiso con el Open Access. En caso de que la institución cuente con una Política sobre Acceso Abierto aparecerá enlazada desde el Repositorio.

Resultado:

- Sí (añadir url)
- No

3.- ASPECTOS LEGALES

En esta sección se recogen aquellas cuestiones relacionadas con la propiedad intelectual de los contenidos distribuidos en el repositorio. Se consideran básicos dos aspectos: primero, garantizar que dichos contenidos, cuando sean archivados por los autores, están libres de restricciones de derechos de copia. Segundo, que el repositorio obtenga del autor, mediante algún tipo de autorización, el derecho no exclusivo de distribución de los contenidos.

3.1.- El autor reconoce que al depositar no está infringiendo ningún derecho de propiedad intelectual. El repositorio facilitará al autor el cumplimiento con la normativa vigente.

En la ingesta de materiales se debe obtener la declaración del autor de que ha respetado los derechos de propiedad intelectual a terceros. Sería, por ejemplo, el caso de las tesis doctorales, etc.

Resultado:

- Sí (especificar cómo y para qué tipo de materiales)
- No

3.2.- Existe una autorización por el autor o el titular de los derechos que permite la distribución de contenidos.

Para cada documento se debe obtener el permiso del autor o del titular de los derechos de explotación para difundirlo a través del repositorio en las condiciones preestablecidas (licencias Creative Commons, contratos de edición, autorizaciones, etc.). Este permiso puede ser individual (por ejemplo, el depósito de una comunicación a un congreso) o puede darse de forma colectiva para un grupo de documentos (por ejemplo, las revistas que asignan una licencia Creative Commons a todos sus artículos).

Resultado:

- Sí
- No



3.3.- Existe documentación disponible que ayude al autor a decidir si puede o no archivar el documento.

Por ejemplo: un enlace a SHERPA/Romeo y/o DULCINEA

Resultado:

- Sí (especificar / añadir url)
- No

3.4.- Inclusión de la información sobre los derechos de autor en los metadatos exportados por el repositorio.

Se valora en este punto que los metadatos en Dublin Core exportados por el repositorio deben incluir definido y completado el campo rights con todas las declaraciones de administración de derechos para acceder o utilizar el objeto, o una referencia a un servicio que proporcione esta información. (Véase además 4.8 y 4.9)

Resultado:

- Sí
- No

4.- METADATOS

En esta sección se marcan y se definen las características y el formato de los metadatos que deben tener los documentos del repositorio. Para ser incluido en la plataforma RECOLECTA, algunas de las reglas de metadatos que se presentan a continuación son de obligado cumplimiento y otras tienen carácter recomendado. En el anexo 5 se pueden ver todos los criterios de contenido que establece el validador RECOLECTA a los repositorios que quieren ingresar en la plataforma y su nivel de obligatoriedad.

4.1. Se utiliza el formato de metadatos OAI_DC

El objetivo principal de utilizar un formato de metadatos común es facilitar la interoperabilidad.

Resultado:

- Sí
- No

4.2.- Todos los registros contienen el campo título (dc:title).

Texto libre que recoge el nombre por el que el recurso es conocido formalmente.

Se debe conservar el nombre original, el orden y la ortografía del título del recurso. Utilizar mayúsculas únicamente para nombres propios. Los subtítulos deben separarse del título mediante dos puntos.

Resultado:

- Sí
- No



4.3.- Todos los registros contienen el campo descripción (dc:description).

Se debe incluir un resumen (abstract) de la publicación, aunque se puede ofrecer más información siempre y cuando no se utilice para indicar información que corresponda a otros campos.

Resultado:

- Sí
- No

4.4.- Todos los registros contienen el campo tipo de publicación (dc:type).

Tipo de resultado científico del cual el recurso es una manifestación. Describe el tipo de diseminación o el tipo de contenido intelectual del recurso. Se utiliza para explicar al usuario qué tipo de recurso está observando.

Resultado:

- Sí
- No

4.5.- El campo tipo de publicación (dc:type) se asigna según los tipos de documentos OpenAIRE.

El tipo de resultado científico debe estar basado en el vocabulario de tipos OpenAIRE. Ver documento anexo 1.

Resultado:

- Sí
- No

4.6.- El campo tipo de publicación (dc:type) se encuentra conforme a vocabulario establecido por OpenAIRE indicando la versión del documento.

Se debe indicar el estado de validación de la publicación. Ver documento anexo 2.

Resultado:

- Sí
- No

4.7.- Todos los registros contienen un campo de fecha de publicación (dc:date).

Este elemento se asociará a la publicación del recurso.

Resultado:

- Sí
- No

4.8.- El campo fecha (dc:date) se encuentra conforme al formato establecido.

La práctica recomendada para codificar el valor de la fecha se define en el perfil de la norma ISO 8601 y sigue el formato AAAA-MM-DD, donde MM y DD son opcionales.

Resultado:

- Sí
- No



4.9.- Todos los registros contienen un campo de derechos de autor (dc:rights).

Información acerca de los derechos contenidos en y sobre el recurso. Normalmente, un elemento Rights contendrá una declaración de gestión de derechos para acceder o utilizar el objeto o una referencia a un servicio que proporcione dicha información. La información de derechos suele englobar los derechos de propiedad intelectual, copyright y otros derechos relacionados con la propiedad.

Es preferible hacer referencia a un servicio de derechos donde los derechos de reutilización se explican al usuario final mediante una dirección URL. Por ejemplo, la organización Creative Commons.

Resultado:

- Sí
- No

4.10.- Todos los registros contienen información del nivel de acceso (dc:rights).

Información acerca de los niveles de acceso basado en el vocabulario de tipos OpenAire. Ver documento anexo 4.

Resultado:

- Sí
- No

4.11.- Todos los registros contienen el campo autor (dc:creator).

La entidad o persona física responsable principal de crear el contenido del recurso.

Resultado:

- Sí
- No

4.12.- Todos los registros contienen un campo de formato (dc:format) conforme al vocabulario establecido.

La manifestación digital del recurso. Se utiliza la lista registrada de IANA de tipos de medios de Internet (tipos MIME) para seleccionar un término. Ver documento anexo 3.

Resultado:

- Sí
- No

4.13.- Todos los registros contienen un campo de idioma (dc:language).

Idioma del contenido intelectual del recurso.

Resultado:

- Sí
- No



4.14.- El campo de idioma (dc:language) se encuentra conforme al vocabulario establecido.

Para este campo se establece como vocabulario: ISO 639-1, ISO 639-2 o ISO 639-3.

Resultado:

- Sí
- No

4.15.- Todos los registros contienen un campo de identificador (dc:identifier).

La práctica recomendada consiste en identificar el recurso por medio de una cadena o número conforme a un sistema de identificación formal. Entre los ejemplos de sistemas de identificación formal se incluyen el identificador uniforme de recurso (URI, Uniform Resource Identifier), incluido el localizador uniforme de recurso (URL, Uniform Resource Locator), el identificador de objetos digitales (DOI, Digital Object Identifier) y el Handle.

Resultado:

- Sí
- No

4.16.- Existe una política de indización conocida por los autores donde se establezca: lengua, lenguajes utilizados, etc.

A la hora de asignar materias o palabras clave a los documentos, tanto por los autores como por el propio repositorio, existe un documento donde se establezcan los requisitos mínimos a seguir que permitan conseguir una homogeneidad de los contenidos. Tales requisitos pueden ser: la lengua en que deben estar redactadas las entradas o los lenguajes documentales utilizados en el caso de una indización controlada.

Resultado:

- Sí (especificar / añadir url)
- No

4.17.- Se aplica algún sistema de clasificación normalizado.

Disponer de uno o varios sistemas de clasificación normalizados tales como CDU, JEL, UNESCO, etc. es de gran ayuda para realizar recolecciones selectivas por los agregadores y pueden facilitar enormemente la creación de servicios de valor añadido.

Resultado:

- Sí (especificar / añadir url)
- No

4.18.- Se permite la exportación de metadatos en algún otro formato aparte del Dublin Core Simple.

Resultado:

- Sí (especificar)
- No



4.19.- Se utiliza algún formato de metadatos técnicos y/o de conservación.

Resultado:

- Sí (especificar)
- No

5.- INTEROPERABILIDAD

5.1.- Se proveen los datos a través del protocolo OAI-PMH.

Resultado:

- Sí
- No

5.2.- Se identifican los recursos de investigación a través de uno o varios sets.

Resultado:

- Sí
- No

5.3.- Se identifican las publicaciones financiadas por la Unión Europea o por otros financiadores a través de un set denominado 'openaire'.

Resultado:

- Sí
- No

5.4.- Se marcan los registros eliminados.

Los registros eliminados deben marcarse durante, al menos, un periodo de tiempo suficiente, de tal forma que los recolectores puedan identificarlos y eliminarlos de sus bases de datos. De otra forma se corre el riesgo de que registros eliminados del proveedor de datos sigan existiendo en los recolectores.

Resultado:

- Sí
- No

5.5.- El tiempo de vida del testigo de reanudación es de un mínimo de veinticuatro horas.

Los testigos de reanudación se utilizan en respuestas incompletas del servidor OAI-PMH para que el cliente pueda retomar la descarga en un momento posterior. La definición de hasta cuándo se puede retomar debe definirla cada repositorio pero no podrá ser en ningún caso inferior a un día.

Resultado:

- Sí
- No



5.6.- El correo electrónico del administrador del repositorio está disponible en la etiqueta AdminEmail dentro de la respuesta a una orden Identify.

Resultado:

- Sí
- No

5.7.- Existe una declaración de Description en la respuesta a una orden Identify.

Esta etiqueta se utiliza para describir el repositorio de una forma entendible por los recolectores. Existen diferentes esquemas ya definidos en el documento Implementation Guidelines for the Open Archives Initiative Protocol for Metadata Harvesting o bien cada repositorio puede definir el suyo.

Resultado:

- Sí
- No

5.8.- La entrega de registros a través del protocolo OAI-PMH es progresiva a través de lotes.

Resultado:

- Sí
- No

5.9.- El tamaño de los lotes para la entrega de registros está dentro del rango de 100-500 registros.

Está comprobado en la práctica que un número de elementos entre este rango agiliza los procesos de recolección y evita sobrecargas en los repositorios.

Resultado:

- Sí
- No

5.10.- Coincide el formato de la fecha expresado en la orden Identify con el campo datestamp de los registros.

Se debe usar el mismo formato para expresar la fecha de los registros que la definida en la granularidad de la página de identificación (verb Identify). La recolección gradual no es posible cuando la granularidad de entrada no puede ser procesada. El patrón definido en el campo 'granularity' debe coincidir con el valor del campo 'datestamp' de los registros.

Resultado:

- Sí
- No



6.- LOGS Y ESTADÍSTICAS

6.1.- Los logs del servidor web donde está alojado el repositorio se archivan de forma permanente.

Resultado:

- Sí
- No

6.2.- El repositorio proporciona un servicio de estadísticas sobre el uso de los documentos almacenados.

Se proporcionan de forma pública datos de accesos y descargas de forma individualizada para cada documento almacenado.

Resultado:

- Sí (especificar / añadir url)
- No

6.3.- Se realiza un filtrado de accesos de los robots o motores de búsqueda.

Resultado:

- Sí
- No

6.4.- Se realiza un filtrado de doble clics.

Resultado:

- Sí
- No

7.- SEGURIDAD, AUTENTICIDAD E INTEGRIDAD DE LOS DATOS

7.1.- Existe un procedimiento establecido sobre la elaboración de copias de seguridad, tanto del software sobre el que funciona el repositorio, los metadatos y los documentos propiamente dichos.

Resultado:

- Sí (especificar / añadir url)
- No

7.2.- El repositorio utiliza identificadores persistentes para sus contenidos (DOIs, Handles, URNs, etc.).

Resultado:

- Sí (especificar / añadir url)
- No



6.- GLOSARIO ⁶

- **Acceso Abierto, *Open Access*:** Acceso abierto significa la disponibilidad libre en Internet de la literatura de carácter académico o científico, permitiendo a cualquier usuario leer, descargar, copiar, distribuir, imprimir, buscar o enlazar los textos completos de los artículos científicos, y usarlos con cualquier otro propósito legítimo, sin otras barreras financieras, legales o técnicas más que suponga Internet en sí misma. (Definición de Budapest Open Access Initiative (BOAI).
- **Agrupaciones, *sets*:** los sets (agrupaciones) son un componente estándar del protocolo OAI-PMH y se utilizan para acotar (filtrar) partes concretas de un repositorio. Si el repositorio también contiene elementos no textuales, no digitales, elementos de acceso de pago o únicamente elementos de metadatos, puede utilizar el mecanismo de sets para filtrar los elementos al suministrar el contenido a los proveedores de servicio.
- **Archivo, *Archive*:** el término *Archive* en el nombre *Open Archives Initiative* refleja el origen de OAI – la comunidad de los e-prints, en que el término “archivo” es usado como sinónimo de repositorio de artículos científicos.
- **Autoarchivo, *Self-Archive*:** depósito de un documento digital por su autor en repositorio accesible públicamente. Cuando un documento es auto-archivado en un sistema que cumple con los protocolos OAI, puede ser buscado y acceder a él a partir de motores de búsqueda genéricos o específicos, potenciando su impacto.
- **Budapest Open Access Initiative (BOAI):** tal vez la más importante iniciativa del movimiento de Acceso Libre al Conocimiento, que resultó de la reunión que tuvo lugar en Budapest en diciembre de 2001, promovida por el *Open Society Institute* (OSI). La declaración allí aprobada estableció el significado y el ámbito del Acceso Abierto y definió dos estrategias complementarias para promoverlo y alcanzarlo. La vía óptima o dorada y la buena o verde.

Más información: <http://www.budapestopenaccessinitiative.org/boai-10-translations/spanish>

- **Declaración de Berlín, *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*:** esta declaración fue suscrita el 22 de octubre de 2003 por representantes de algunas de las más importantes instituciones científicas europeas, entre ellas la Sociedad Max-Planck (Alemania) o el *Centre National de la Recherche Scientifique* (Francia), apoyando el *Open Access* y el depósito en repositorios de acceso abierto, y afirmando animar a sus investigadores y científicos a depositar sus trabajos en, por lo menos, un repositorio.

Más información: <http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html>

⁶ (*) Las definiciones marcadas con (*) provienen de del glosario de términos elaborados por REBIUN para la encuesta de “Políticas de acceso abierto” desarrollada por el grupo de su Línea Estratégica 2.



- **Derechos de Autor, copyright:** los derechos de autor, también llamados copyright en las tradiciones jurídicas anglosajonas, pertenecen al creador intelectual de la obra, salvo disposición expresa en contrario. Se consideran como obras las creaciones originales literarias, artísticas o científicas expresadas por cualquier medio o soporte, tangible o intangible, actualmente conocido o que se invente en el futuro, y que, como tales, son protegidas en los términos de la legislación española en materia de propiedad intelectual (Ley 23/2006, de 7 de julio, por la que se modifica el texto refundido de la Ley de Propiedad Intelectual, aprobado por el Real Decreto Legislativo 1/1996, de 12 de abril).

Los derechos de autor son de dos tipos:

- los derechos morales de autoría e integridad de la obra, que son irrenunciables e inalienables.
 - los derechos de explotación de su obra en cualquier forma y, en especial, los derechos de reproducción, distribución, comunicación pública y transformación, que no podrán ser realizadas sin el permiso del autor.
- **DINI (*Deutsche Initiative für Netzwerkinformation E.V.*):** modelo para asegurar la calidad de los repositorios digitales en Alemania. DINI fue creada para promover la mejora de los servicios de información y de las comunicaciones, así como el desarrollo de las infraestructuras de la información. Pretende mejorar de forma continua la oferta de tecnología de la información y de servicios. Más información: <http://edoc.hu-berlin.de/series/dini-schriften/2010-3-en/PDF/dini-zertifikat-2010-3-en.pdf>
 - **Documentación Institucional** (*): documentos generados por una institución y que no son producto de la actividad docente e investigadora de la misma.
 - **DRAMBORA (*Digital Repository Audit Method Based on Risk Assessment*):** herramienta que sirve para analizar un repositorio digital, determinando sus capacidades, debilidades y fortalezas.
Más información: <http://www.dcc.ac.uk/resources/repository-audit-and-assessment/drambora>
 - **DRIVER (*Digital Repository Infrastructure Vision for European Research*):** proyecto europeo para construir una infraestructura pública que facilite el acceso abierto a la información. La idea del proyecto es que cualquier información científica, técnica, datos experimentales y de observaciones debe estar disponible en Internet y de forma gratuita. Más información: <http://www.driver-repository.eu/Driver-About/About-DRIVER.html>
 - **Dublin Core:** el *Dublin Core* surgió en el encuentro realizado en Dublin (Ohio), en 1995, y fue concebido para describir recursos de Internet y responder a la necesidad de crear una nomenclatura común de metadatos. Está compuesto de quince elementos de descripción de documentos electrónicos - title, creator, subject, description, publisher, contributor, date, type, format, identifier, language, source, relation, coverage, rights – que facilitan la información básica sobre los mismos. El Dublin Core Cualificado es una extensión del Dublin Core donde algunos de sus elementos son acompañados de un cualificador que los hace más restrictivos. Más información: <http://dublincore.org/>



- **Dulcinea:** DULCINEA (<http://www.accesoabierto.net/dulcinea/>) , por su analogía con el proyecto ROMEO Sherpa (<http://www.sherpa.ac.uk/about.html>), es un proyecto cuyo objetivo es conocer las políticas editoriales de las revistas españolas respecto al acceso a sus archivos, los derechos de copyright sobre los mismos y cómo estos pueden afectar a su posterior autoarchivo en repositorios institucionales o temáticos.
- **E-print:** versión digital de un artículo científico con peer-review, antes o después de la evaluación y publicación.
- **Fondo antiguo** (*): Fondo bibliográfico patrimonial de una institución. Es frecuente que sea digitalizado e incluido en el repositorio de una institución.
- **Interoperabilidad** (*): Es la capacidad de los repositorios para comunicarse e intercambiar procesos y datos. Es posible gracias al protocolo OAI-PMH.
- **Licencias Creative Commons:** *Creative Commons* es una organización internacional sin ánimo de lucro, cuya finalidad es la difusión del conocimiento y de la cultura a través de la creación de licencias con las que los creadores pueden indicar qué usos permiten y cuales no en relación a sus obras. Más información: <http://es.creativecommons.org/blog/>
- **Metadatos:** Literalmente, los metadatos son datos sobre datos. Los metadatos son datos asociados a objetos o sistemas de información para fines de descripción, administración, uso, preservación, etc. Existen varios tipos de metadatos: descriptivos (que sirven para identificación y localización), administrativos (creación, derechos, control de acceso, etc.) y estructurales (que relacionan los objetos).
- **Objetos de aprendizaje** (*): materiales digitales elaborados por el profesorado como apoyo al aprendizaje de sus alumnos.
- **OpenAIRE:** infraestructura de acceso abierto a la investigación en Europa, creada y desarrollada dentro de los proyectos del 7^º Programa Marco OpenAIRE y OpenAIRE Plus desde el año 2009. OpenAIRE tiene el recolector (www.openaire.eu) de acceso abierto en Europa, que recolecta las publicaciones desarrolladas en el marco de los proyectos de investigación afectados por los requisitos de acceso abierto de la Comisión Europea. El mecanismo de recolección de OpenAIRE se basa en el protocolo OAI-PMH. Tecnológicamente OpenAIRE utiliza D-NET para recolectar los metadatos como motor de búsqueda de D-NET. Este motor es el encargado de recolectar los “sets OpenAIRE” de la red de repositorios institucionales registrados y auto-validados previamente como “OpenAIRE Compatible”. Posteriormente a su recolección, OpenAIRE realiza un enriquecimiento de datos en base a la información de que disponen sobre los proyectos subvencionados en el 7PM. Este enriquecimiento de datos se realiza incluyendo información relevante del proyecto (identificador del proyecto, identificador del proyecto 7PM, fecha de publicación, fecha de inicio, fecha de fin, etc.). Para ofrecer los servicios de búsqueda, el buscador D-NET se integra en el portal OpenAIRE, siendo transparente para los usuarios.



- **Open Archives Initiative (OAI):** surge en 1999 con el objetivo de crear una plataforma simple para permitir la interoperabilidad y la búsqueda de publicaciones científicas de diversas disciplinas. Esta iniciativa se enclava en el seno de la comunidad de los “e-prints” y partió de un inicio esencialmente técnico (de lo que resultó el protocolo OAI-PMH: <http://www.openarchives.org/pmh/>). Proporciona una base estable para la interoperabilidad de archivos abiertos, lo que hace que crezca el número de servidores que lo implementan, contribuyendo a dar mayor visibilidad y difusión al movimiento de Acceso Libre al Conocimiento.
- **Peer-review, revisión por pares; revisión científica:** es el proceso de evaluación y certificación de calidad de la investigación y de sus resultados, realizado en el momento de su publicación. Los artículos de las principales revistas científicas son objeto de este proceso antes de su publicación.
- **Política institucional de acceso abierto** (*): conjunto de acciones que se compromete a llevar a cabo una institución para fomentar e implantar el acceso abierto a su producción científica. En el caso de las universidades es aprobada por el Consejo de Gobierno.
- **Post-print:** el texto digital de un artículo que ha sido evaluado y revisado y que ha sido aceptado para su publicación por una revista científica.
- **Pre-print:** el texto digital de un artículo que aún no ha sido evaluado y revisado y aún no ha sido aceptado para su publicación por una revista científica.
- **Recolector de recursos digitales** (*): plataforma en acceso abierto que periódicamente recopila los metadatos de los repositorios que colaboran con ella. Facilitan el acceso a contenidos científicos procedentes de múltiples repositorios, contribuyendo notablemente a su visibilidad. La recolección de metadatos es posible si los repositorios cumplen el protocolo OAI-PMH.
- **Recomendación** (*): Política institucional a favor del acceso abierto mediante la cual se aconseja e insta, pero no se obliga, a los autores a depositar el resultado de su producción académica y científica en el repositorio de la institución.
- **Repositorio / Repositorio Institucional** (*): es un conjunto de servicios que ofrece una institución a los miembros de su comunidad para la gestión y distribución de materiales digitales creados por la institución y sus miembros, organizados de tal modo que se garantice la correcta administración de los mismos, incluyendo la preservación a largo plazo, el acceso y su difusión. Otras denominaciones por las que puede ser conocido: depósito digital institucional; archivo digital; repositorio digital.
- **Sherpa/Romeo:** es un proyecto cuyo objetivo es conocer las políticas editoriales de las revistas científicas respecto al acceso a sus archivos, los derechos de copyright sobre los mismos y cómo estos pueden afectar a su posterior autoarchivo en repositorios institucionales o temáticos.

Más información: <http://www.sherpa.ac.uk/romeo/index.php?fidnum=|&mode=simple&la=es>



ANEXO 1 VOCABULARIO ESTABLECIDO PARA EL CAMPO TIPO DE PUBLICACIÓN (DC:TYPE)

TÉRMINO	DESCRIPCIÓN
article (Artículo científico antes de ser publicado, versión del editor)	Un artículo científico (a veces también llamado <i>paper</i>) es un trabajo relativamente breve destinado a la publicación en revistas especializadas. Tiene como objetivo difundir de manera clara y precisa, los resultados de una investigación realizada sobre un área determinada del conocimiento. También busca fomentar el desarrollo de métodos experimentales innovadores. Debe estar cuidadosamente redactado para expresar de un modo claro y sintético lo que se pretende comunicar, y para que incluya las citas y referencias, indispensables para verificar y poder reproducir los resultados originales que se dan a conocer en el mismo.
bachelorThesis (Proyecto fin de carrera. Trabajo final de grado)	Son proyectos exigidos al final de ciertas carreras académicas como condición para obtener la titulación. Este proyecto suele tener una duración de alrededor de un año. Con este trabajo se adquiere el título de Grado. También denominado TFG.
masterThesis (Tesina. Trabajo final de máster o postgrado)	Es un trabajo de investigación relacionado con los estudios cursados en un máster. Con este trabajo con el que se adquiere el grado de Máster. También denominado TFM.
doctoralThesis (Tesis doctoral)	Es la tesis de más alto nivel. Es un estudio escrito tras cuatro o cinco años de investigación sobre un tema dado. La tesis doctoral es un trabajo inédito de investigación que permite acceder al grado de Doctor.
book (Libro o monografía)	Es un documento que trata un tema en particular. Utiliza diversas fuentes compiladas y procesadas por uno o por varios autores.
bookPart (Capítulo o parte de un libro)	Es una de las principales divisiones de una obra escrita de cierta longitud como un libro.
review (Revisión, evaluación de un libro o artículo)	Es una evaluación de un artículo o de un libro en el que se analiza el contenido, el estilo y el mérito del autor.
conferenceObject (Publicaciones de conferencias: comunicaciones, ponencias, pósters, etc)	Se incluyen todo tipo de documentos relacionados con una conferencia. Por ejemplo: comunicaciones a congresos, informes de conferencias, ponencias, trabajos publicados en actas de congresos (<i>proceedings</i>), contribuciones a congresos, resúmenes de conferencias (<i>abstracts of conference papers</i>), posters.
lecture (Conferencia académica)	Es una presentación en un acto académico. Por ejemplo una conferencia inaugural.
workingPaper (Artículo preliminar)	Es un artículo científico o técnico preliminar que se publica en la institución donde se realiza la investigación. También se conoce como trabajo de investigación, exposición de la investigación o documento de debate (<i>research paper, research memorandum or discussion paper</i>). La diferencia entre un <i>preprint</i> y un <i>workingPaper</i> es que se comparte/publica en la institución.



TÉRMINO	DESCRIPCIÓN
<p>preprint (Artículo científico antes de ser publicado, versión del autor)</p>	<p>Es la denominación de un artículo o trabajo científico en versión previa a su publicación y antes de la revisión por pares (<i>peer-review</i>). A diferencia del <i>workingPaper</i> este artículo científico o técnico preliminar no se ha publicado en la institución. El <i>preprint</i> está destinado a ser publicado en una revista científica o como un capítulo de un libro.</p>
<p>report (Estudio, informe, memoria)</p>	<p>En esta categoría se incluyen informes de comisiones, memorandos, informes de investigación externos, informes internos, informes estadísticos, informes a una agencia de financiación, documentación técnica, entregables de un proyecto, etc.</p>
<p>annotation (Nota a un texto de jurisprudencia)</p>	<p>Es una nota a un texto legal.</p>
<p>contributionToPeriodical (artículo de prensa)</p>	<p>Es un artículo para un periódico, una revista semanal u otra revista periódica no académica.</p>
<p>patent (patente)</p>	<p>Es un conjunto de derechos exclusivos concedidos por un Estado al inventor o inventores de un nuevo producto susceptible de ser explotado industrialmente, por un período limitado de tiempo a cambio de la divulgación de la invención.</p>
<p>other (otros)</p>	<p>Otros documentos como datos de investigación no publicados (<i>non-publication research data</i>), materiales audiovisuales, animaciones, etc.</p>

Fuente: https://guidelines.openaire.eu/wiki/Literature_Guidelines:_Metadata_Field_Publication_Type



ANEXO 2 VOCABULARIO ESTABLECIDO PARA INDICAR LA VERSIÓN DEL DOCUMENTO EN EL CAMPO TIPO DE PUBLICACIÓN (DC:TYPE)

TÉRMINO	DESCRIPCIÓN
draft (borrador)	Versión temprana de un trabajo aún en desarrollo
submittedVersion (versión enviada)	Versión que se ha presentado a una revista para la revisión por pares
acceptedVersion (versión aceptada)	Versión del autor que incorpora los comentarios de los evaluadores y es la versión aceptada para publicación
publishedVersion (versión publicada)	Versión del editor
updatedVersion (versión actualizada)	Versión actualizada desde la publicación

Fuente: https://guidelines.openaire.eu/wiki/Literature_Guidelines:_Metadata_Field_Publication_Version



ANEXO 3 VOCABULARIO ESTABLECIDO PARA EL CAMPO TIPO DE FORMATO (DC:FORMAT)

Se utiliza la lista registrada de IANA de tipos de medios de Internet (tipos MIME) para seleccionar un término. <http://www.iana.org/assignments/media-types/>

Las principales son las siguientes:

TIPO	SUBTIPO
Text (texto)	<ul style="list-style-type: none"> • plain • richtext • enriched • tab-separated-values • html • sgml • xml
Application (aplicación)	<ul style="list-style-type: none"> • octet-stream • postscript • rtf • applefile • mac-binhex40 • wordperfect5.1 • pdf • vnd.oasis.opendocument.text • zip • macwriteii • msword • sgml • ms-excel • ms-powerpoint • ms-project • ms-works • xhtml+xml • xml
Image (imagen)	<ul style="list-style-type: none"> • jpeg • gif • tiff • png • jpeg2000 • sid
Audio (audio)	<ul style="list-style-type: none"> • wav • mp3 • quicktime
Video (video)	<ul style="list-style-type: none"> • mpeg1 • mpeg2 • mpeg3 • avi



ANEXO 4 VOCABULARIO ESTABLECIDO PARA EL CAMPO NIVEL DE ACCESO (DC:RIGHTS)

TÉRMINO	DESCRIPCIÓN
closedAccess (acceso cerrado)	Este tipo de acceso es opuesto al acceso abierto.
embargoedAccess (embargado)	El recurso es de acceso cerrado, hasta que sea liberado en acceso abierto en una fecha determinada (esta fecha puede especificarse en otro apartado de los metadatos).
openAccess (acceso abierto)	Acceso sin restricciones al recurso.
restrictedAccess (restringido)	Acceso abierto con restricciones. Con frecuencia los autores quieren controlar a las personas que tienen acceso y facilitan el acceso a sus obras a través de una dirección de correo electrónico, para mantenerse en contacto con los que están interesados.

Fuente: https://guidelines.openaire.eu/wiki/Literature_Guidelines:_Metadata_Field_Access_Level



ANEXO 5 CRITERIOS RECOLECTA PARA CONTENIDO

Este anexo recoge las reglas de metadatos que el validador RECOLECTA analiza, marcando claramente cuáles son de obligado cumplimiento.

REGLAS DE METADATOS			
METADATO	REGLA	VALIDADOR RECOLECTA	CRITERIO
TITLE	R1: Todos los registros contienen el campo título (dc:title).	obligatorio	4.2
DESCRIPTION	R3: Todos los registros contienen el campo descripción (dc:description).	recomendado	4.3
TYPE	R5: Todos los registros contienen el campo tipo de publicación (dc:type).	obligatorio	4.4
	R6: El campo dc:type se encuentra conforme al vocabulario establecido.	obligatorio	4.5
	R7: El campo tipo de publicación (dc:type) se encuentra conforme a vocabulario establecido indicando la versión del documento.	recomendado	4.6
DATE	R8: Todos los registros contienen un campo de fecha de publicación (dc:date).	obligatorio	4.7
	R10: El campo dc:date se encuentra conforme al formato establecido.	obligatorio	4.8
RIGHTS	R11: Todos los registros contienen un campo de derechos de autor (dc:rights).	obligatorio	4.9
	R12: El campo derechos de autor (dc:rights) contiene información del nivel de acceso. Éste se encuentra conforme al vocabulario establecido.	recomendado	4.10
CREATOR	R13: Todos los registros contienen el campo autor (dc:creator).	obligatorio	4.11
FORMAT	R15: Todos los registros contienen un campo de formato (dc:format) conforme al vocabulario establecido.	obligatorio	4.12
LANGUAGE	R16: Todos los registros contienen un campo de idioma (dc:language).	obligatorio	4.13
	R17: El campo de idioma (dc:language) se encuentra conforme al vocabulario establecido.	obligatorio	4.14
IDENTIFIER	R18: Todos los registros contienen un campo de identificador (dc:identifier).	obligatorio	4.15



ANEXO 6 CRITERIOS RECOLECTA OBLIGATORIOS PARA PROTOCOLO OAI-PMH

REGLAS DE PROTOCOLO OAI-PMH PARA UNA CORRECTA RECOLECCIÓN DEL RECURSO
R1: Es obligatorio utilizar el formato de metadatos OAI_DC
R2: El tamaño del lote debe estar dentro del rango de 100-500 registros
R3: Implementación de una estrategia de eliminación: transitoria o persistente
R4: Coinciden los formatos de las fechas
R5: Entrega progresiva de registros
R6: Tiempo de vida mínimo para el ' <i>ResumptionToken</i> ' debe ser de 24 horas. El validador no analizará esta regla para evitar que falle cuando el ' <i>ResumptionToken</i> ' tenga una duración ilimitada.
R7: Dirección de correo electrónico del administrador del repositorio válida



BIBLIOGRAFÍA

- Crow, R. (2002): *SPARC Institutional Repository Checklist & Resource Guide*. Washington, DC: Scholarly Publishing & Academic Resources Coalition. Versión 1.0. November 2002. http://sparc.arl.org/sites/default/files/presentation_files/ir_guide_checklist_v1.pdf [Consulta 4/10/2013]
- Deutsche Initiative für Netzwerkinformation (2011). *DINI-Certificate*. [Version 3.0, March 2011] <http://edoc.hu-berlin.de/series/dini-schriften/2010-3-en/PDF/dini-zertifikat-2010-3-en.pdf> [Consulta 4/10/2013]
- DigitalPreservationEurope (2008). *Repository Planning Checklist and Guidance PLATTER*. www.digitalpreservationeurope.eu/platter.pdf [Consulta 4/10/2013]
- Dobratz, S. y F. Scholze (2006): "DINI Institutional Repository Certification and Beyond." *Library Hi Tech*, vol. 24, núm. 4 (2006), págs. 583-594.
- DRIVER Project (2008): *Directrices DRIVER 2.0. Directrices para proveedores de contenido - Exposición de recursos textuales con el protocolo OAI-PMH*. http://recolecta.fecyt.es/sites/default/files/contenido/documentos/DRIVER_2_1_Guidelines_Spanish.pdf [Consulta 4/10/2013]
- European Commission (2013): *Fact Sheet: Open Access is Horizon 2020*. https://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/FactSheet_Open_Access.pdf [Consulta 29/01/2014]
- _____ (2013): *Commission launches pilot to open up publicly funded research data*. Nota de prensa 16 Diciembre 2013. http://europa.eu/rapid/press-release_IP-13-1257_en.htm [Consulta 29/01/2014]
- _____ (2013): *Guidelines on Data Management in Horizon 2020*. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf [Consulta 29/01/2014]
- _____ (2013): *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020*. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf [Consulta 29/01/2014]
- Melero, R., E. Abadal, F. Abad y J.M. Rodríguez-Gairín (2009): *Situación de los repositorios institucionales en España: informe 2009*. http://www.accesoabierto.net/sites/accesoabierto.net/files/Informe2009-Repositorios_0.pdf [Consulta 4/10/2013]
- OpenAIRE (2010). *Directrices OpenAIRE 1.1: Directrices para proveedores de contenido del espacio de información OpenAIRE*. http://www.openaire.eu/index.php?option=com_content&view=article&id=207&lang=es [Consulta: 4/10/2013]
- _____ (2013): *OpenAIRE Guidelines wiki*. <http://guidelines.openaire.eu> [Consulta: 17-03-2014]

REBIUN (2009): *Informe de las actuaciones realizadas en el desarrollo del objetivo operacional 1.1.1: consolidar la implementación del portal Recolecta en el marco del proyecto europeo Driver*. XVII Asamblea Anual de REBIUN, León, 4-6 de noviembre 2009 http://www.rebiun.org/documentos/Documents/IIPE_LINEA2_07-11/IIPE_Linea2_Guia_evaluacion_repositorios_2010.pdf [Consulta 14/05/2010]

_____ (2012): *Informe de resultados de la encuesta de “Políticas de acceso abierto”*. Miembros del Grupo de la Línea 2 objetivo 4 de REBIUN. 17 a 31 de octubre de 2012.

The Center for Research Libraries (OCLC) and Online Computer Library Center (OCLC) (2007): *Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC)*. Versión 1.0, febrero 2007. www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf [Consulta 4/10/2013]

VVAA (2008): *Directrices para la creación de repositorios institucionales en universidades y organizaciones de educación superior*. Pontificia Universidad Católica de Valparaíso. http://eprints.rclis.org/13512/1/Directrices_RI_Espa_ol.pdf [Consulta 4/10/2013]

WEBGRAFÍA

BuscaRepositorios. Directorio de repositorios institucionales españoles de acceso abierto. <http://www.accesoabierto.net/repositorios/> [Consulta: 4/10/2013]

Digital Repository Audit Method Based on Risk Assessment (DRAMBORA). <http://www.repositoryaudit.eu/> [Consulta 4/10/2013]

NESTOR Catalogue of Criteria for Trusted Digital Repositories. <http://www.dcc.ac.uk/tools/neighbor/> [Consulta 4/10/2013]

OpenDOAR. Directory of Open Access Repositories. <http://www.opendoar.org> [Consulta: 4/10/2013]

Proyecto PLANETS: Preservation and Long-term Access through NETworked Services. <http://www.planets-project.eu/> [Consulta 4/10/2013]

Ranking web de repositorios del mundo. <http://repositories.webometrics.info/es> [Consulta 4/10/2013]

RECOLECTA: Recolector de ciencia abierta. <http://recolecta.fecyt.es/> [Consulta 4/10/2013]