

Programa de Doctorado Formación en la Sociedad del Conocimiento

**Visibilidad e impacto de la literatura gris científica en repositorios institucionales
de acceso abierto. Estudio de caso bibliométrico del repositorio Gredos de la
Universidad de Salamanca**

TESIS DOCTORAL

Directores:

Dr. D. José Antonio Merlo Vega



Dr. D. Francisco José García-Peñalvo



Doctoranda:

M^a del Tránsito Ferreras Fernández



Octubre, 2016



GRupo de Investigación en
InterAcción y eLearning (GRIAL)

Dr. José Antonio Merlo Vega, profesor Titular del Departamento de Biblioteconomía y Documentación de la Universidad de Salamanca, y Dr. Francisco José García Peñalvo, profesor Titular del Departamento de Informática y Automática de la Universidad de Salamanca, en calidad de directores del trabajo de tesis doctoral titulado “Visibilidad e impacto de la literatura gris científica en repositorios institucionales de acceso abierto. Estudio de caso bibliométrico del repositorio Gredos de la Universidad de Salamanca” y realizado por Dña. M^a del Tránsito Ferreras Fernández,

HACEN CONSTAR

Que dicho trabajo tiene suficientes méritos teóricos contrastados adecuadamente mediante las validaciones oportunas, publicaciones relacionadas y aportaciones novedosas. Por todo ello manifiestan su acuerdo para que sea autorizada la presentación y defensa del trabajo referido.

En Salamanca, a 26 de octubre de 2016,

Fdo. José Antonio Merlo Vega



Fdo. Francisco José García Peñalvo

ÍNDICE DE CONTENIDO

Índice de contenido.....	i
Agradecimientos	ix
Lista de abreviaturas	xiii
Resumen	xxi
Módulo 0. Introducción	29
<u>Capítulo 0.1. Objeto de estudio</u>	31
<u>Capítulo 0.2. Objetivos</u>	35
<u>Capítulo 0.3. Motivación / Oportunidad de la investigación</u>	37
<u>Capítulo 0.4. Preguntas de investigación</u>	47
<u>Capítulo 0.5. Estudio de la bibliografía</u>	49
<u>Capítulo 0.6. Estructura</u>	53
<u>Capítulo 0.7. Marco teórico</u>	61
<u>Capítulo 0.8. Marco metodológico</u>	67
Módulo I. Estado de la cuestión	77
<u>Capítulo I.1. Revisión sistemática de la literatura</u>	80
<u>I.1.1. Introducción</u>	80
<u>I.1.2. Planteamiento del problema</u>	81
<u>I.1.3. La revisión de la literatura</u>	83
<u>I.1.4. El proceso de revisión</u>	91
<u>Capítulo I.2. El conocimiento científico abierto</u>	127
<u>I.2.1. La comunicación científica</u>	130

<u>I.2.2. El acceso abierto como nuevo modelo de comunicación científica</u>	137
<u>I.2.3. Las ventajas del acceso abierto al conocimiento</u>	144
<u>I.2.4. Contenidos del acceso abierto</u>	150
<u>I.2.5. Políticas de acceso abierto</u>	151
<u>I.2.6. Las diez claves del acceso abierto</u>	162
<u>Capítulo I.3. Repositorios Institucionales (RIs) de acceso abierto</u>	164
<u>I.3.1. Los repositorios institucionales. Definición y datos</u>	166
<u>I.3.2. Uso del repositorio y captación del contenido</u>	174
<u>I.3.3. Factores de éxito de los Repositorios Institucionales</u>	177
<u>I.3.4. La interoperabilidad y visibilidad de los repositorios</u>	182
<u>I.3.5. Difusión, visibilidad, uso e impacto de los contenidos de los repositorios</u>	197
<u>Capítulo I.4. Ecosistemas tecnológicos. El repositorio Gredos</u>	201
<u>I.4.1. Ecosistemas tecnológicos: una aproximación</u>	201
<u>I.4.2. El repositorio Gredos</u>	204
<u>Capítulo I.5. La literatura gris científica. Las tesis doctorales</u>	211
<u>I.5.1. Literatura gris. Definición y datos</u>	211
<u>I.5.2. Desarrollo de políticas para la gestión de Literatura Gris</u>	215
<u>I.5.3. La evaluación de Literatura Gris</u>	216
<u>I.5.4. Las tesis doctorales. Las tesis electrónicas en acceso abierto</u>	218
<u>I.5.5. Beneficios y riesgos de las tesis en acceso abierto: percepción de los autores</u>	225
<u>I.5.6. El aumento de las e-tesis en acceso abierto. El valor de las e-tesis en los RIs</u>	228
<u>I.5.7. Diseminación de las e-tesis. Redes, recolectores y portales de e-tesis</u>	237
<u>I.5.8. Acceso a las e-tesis: mejoras para el acceso</u>	245

<u>Capítulo I.6. Estudios de impacto de la investigación en acceso abierto. Visibilidad, uso y citación de las tesis doctorales</u>	249
<u>I.6.1. Estudios de impacto de la investigación en acceso abierto</u>	250
<u>I.6.2. Visibilidad, uso y citación de las tesis doctorales</u>	259
<u>Capítulo I.7. Bibliometría para el acceso abierto. Métricas alternativas</u>	267
<u>I.7.1. Bibliometría. Definición y terminología</u>	267
<u>I.7.2. La comunicación científica en línea y la bibliometría tradicional</u>	271
<u>I.7.3. Nuevas herramientas y métricas de la ciencia en un escenario de acceso abierto</u>	273
<u>Módulo II. Sistema de análisis y caso de estudio</u>	285
<u>Capítulo II.1. Caso de estudio: Tesis doctorales del repositorio Gredos</u>	287
<u>II.1.1. Colección de Tesis doctorales en el repositorio Gredos</u>	288
<u>II.1.2. Normativa sobre el depósito de tesis en Gredos</u>	289
<u>II.1.3. Depósito de las tesis doctorales en Gredos</u>	291
<u>II.1.4. Diseminación de las tesis doctorales de Gredos: la interoperabilidad</u>	294
<u>Capítulo II.2. Material y métodos</u>	299
<u>II.2.1. Muestra</u>	299
<u>II.2.2. Selección del material</u>	300
<u>II.2.3. Definición de las variables</u>	307
<u>II.2.4. Análisis y métodos estadísticos</u>	309
<u>Módulo III. Resultados y discusión</u>	317
<u>Capítulo III.1. Resultados</u>	319
<u>III.1.1. Resultados del análisis de citación en WOS</u>	319
<u>III.1.2. Resultados del análisis de visibilidad, uso y citación de las tesis de Gredos</u>	320

<u>Capítulo III.2. Discusión</u>	329
<u>III.2.1. Cuestiones previas</u>	329
<u>III. 2.2. WOS como herramienta de búsqueda de tesis citadas</u>	330
<u>III. 2.3. Medición de la visibilidad, el uso y el impacto de las tesis doctorales en OA</u>	332
<u>III.2.4. Google Scholar como herramienta alométrica para la identificación de citas</u>	336
<u>Módulo IV. Conclusiones</u>	341
<u>Capítulo IV.1. Conclusiones</u>	343
<u>Capítulos IV.2. Publicaciones</u>	347
<u>Capítulo IV.3. Investigaciones futuras</u>	351
<u>Módulo V. Bibliografía</u>	355
<u>Módulo VI. Anexos</u>	389
<u>Anexo VI.1. Índice de tablas</u>	391
<u>Anexo VI.2. Índice de figuras</u>	393
<u>Anexo VI.3. Revisión sistemática de la literatura</u>	395
<u>Anexo VI.4. Base de datos de las tesis doctorales analizadas</u>	397

Dedicatoria

A:

Luis, mi marido

y

Alicia, mi hija

Agradecimientos

Esta tesis doctoral es el resultado de cuatro años de trabajo e investigación que solo ha sido posible gracias a la ayuda, ánimo y amistad recibidos en todo momento de mis directores de tesis, el Dr. José Antonio Merlo Vega y el Dr. Francisco José García Peñalvo. Mi agradecimiento más sincero a ambos por compartir su inestimable sabiduría y experiencia, no solo durante el transcurso de la realización de esta tesis, sino durante todos los años de relación laboral e investigadora que la precedieron en el Servicio de Bibliotecas y en el repositorio Gredos.

A la Dra. Helena Martín Rodero, buena compañera de trabajo y mejor amiga, por su generosa ayuda y aportación en los temas de análisis estadístico y revisión sistemática de la literatura. Mi agradecimiento, además, y sobre todo, por su apoyo incondicional y afecto que me ayudaron a sobreponerme en los muchos momentos sufridos de desánimo y de inseguridad.

A mi marido Luis por estar en todo momento a mi lado. Gracias por su comprensión, apoyo y cariño a lo largo de estos años. Sin él, la realización de este proyecto no hubiera sido posible.

Y por último, a mi hija Alicia por ser la razón primera y más importante de mi vida.

Lista de abreviaturas y siglas

APC – Article processing charge

APA – American Psychological Association

API – Application Programming Interface

ARL - Association of Research Libraries

AUP- Amsterdam University Press

AuthorClaim – Aplicación para la identificación de autores

BASE – Bielefeld Academic Search Engine

BBB - Budapest-Bethesda-Berlin

BOAI – Budapest Open Access Initiative

CERIF – Common European Research Information Format

COAR – Confederation of Open Access Repositories

COPE – Committee on Publication Ethics

COUNTER – Counting Online Usage of Networked Electronic Resources

CRIS – Current Research Information Systems

CRIS-OAR – Current Research Information Systems and Open Access Repositories

DART Europe – Portal de acceso a las tesis doctorales de Europa

DataCite – Organización internacional que promueve la citación de los datos

DC – Dublin Core Metadata Standard

DINI – Certificate for Document and Publication Services

DOAJ – Directory of Open Access Journals

DOI – Digital Object Identifier

DRIVER – Digital Repository Infrastructure Vision for European Research

EC – European Commission

ERA – European Research Area

EROMM – European Register of Microform and Digital Masters

ESE – E Europeana Semantic Elements

ETD – Electronic Theses and Dissertations

EUNIS – European University Information Systems

EUROCRIS – European Current Research Information Systems

EUROPEANA – Biblioteca Digital Europea de acceso libre

EUSAL – Ediciones Universidad de Salamanca

FECYT - Fundación Española para la Ciencia y Tecnología

FP7 – Séptimo Programa Marco de la Comisión Europea (2007-2013)

Gredos – Gestión del Repositorio Documental de Salamanca

Handle System – Especificación tecnológica para asignar, gestionar y resolver identificadores persistentes para objetos digitales y recursos en Internet

Hispana – Portal de acceso a la cultura digital y agregador nacional de contenidos a Europeana

HORIZON 2020 – Programa Marco de la Unión Europea para la Investigación y la Innovación (2014-2020)

IR - Institutional Repository

ISBN – International Standard Book Number

ISSN – International Standard Serial Number

I+D – Investigación y Desarrollo

JCR – Journal Citation Reports

JICS - Joint Information Systems Committee

JIF - Journal Impact Factor

KE Usage Statistics Guidelines – Knowledge Exchange and Usage Statistics Guidelines

LG – Literatura gris

LIS – Library and Information Sciences

METS – Metadata Encoding and Transmission Standard

MODS – Metadata Object Description Schema

NDLTD – Networked digital Library of Theses and Dissertations

NIH National Institutes of Health, United States

OA – Open Access

OA-RJ – Open Access Repository Junction

OAI-ORE – Open Archives Initiative – Object Reuse and Exchange

OAI-PMH – Open Archives Initiative – Protocol for Metadata Harvesting

OAIS – Open Archival Information System

OAS - OA-Statistik

OASPA - Open Access Scholarly Publishers Association

OATD – Open Access Theses and Dissertations

OAW – Open Access Week

OJS – Open Journal System

OpenAIRE – Open Access Infrastructure Research for Europe

OpenAIREplus – Segunda generación de OpenAire

OpenDOAR – Directory of Open Access Repositories

ORI – Organisation and Repository Identification

ORCID – Open Researcher & Contributor ID

PersID – Persistent Identifier

PICO – Problem, Intervention, Comparison, Outcome

PICOC – Problem, Intervention, Comparison, Outcome, Context

PIRUS – Publishers and Institutional Repository Usage Statistics

PIRUS2 – Publishers and Institutional Repository Usage Statistics 2

PLoS – Public Library of Science

PMC – PubMed Central

PREMIS – Preservation Metadata: Implementation Strategies

RCUK – Research Councils of UK

RDF – Resource Description Framework

RI – Repositorio Institucional

ROAR – Registry of Open Access Repositories

SEO – Search engine optimization

SJR – SCImago Journal & Country Rank

SURE – Statistics on the Usage of Repositories

SWORD – Simple Web-service Offering Repository Deposit

TDR – Tesis Doctorales en Red

TDX – Tesis Doctorals en Xarxa

TIC – Tecnologías de la Comunicación y la Información

UK RepositoryNet+ (RepNet) – Infraestructura de repositorios Open Access del Reino Unido

UNESCO – United Nations Educational, Scientific and Cultural Organization

URL – Uniform Resource Locators

URN – Uniform Resource Names

XML – Extensible Markup Language

Web – World Wide Web

WOS – Web of Science

WWW – World Wide Web

Resumen

Objetivo

El principal objetivo de este trabajo es demostrar los beneficios del acceso abierto como modelo de comunicación científica para la literatura gris científica en general y para las tesis doctorales electrónicas en particular. El trabajo de investigación que se presenta pretende dar respuesta a la pregunta sobre la visibilidad, usabilidad e impacto de las tesis doctorales depositadas en repositorios de acceso abierto, a través de un estudio de caso bibliométrico del repositorio Gredos de la Universidad de Salamanca.

Método

La presente investigación se compone de dos partes: una teórica y otra práctica. Para la localización y selección de la literatura pertinente en la que sustentar las dos partes de la investigación se llevó a cabo una revisión sistemática de la literatura.

En la parte teórica se realizó un estudio pormenorizado del estado de la cuestión sobre el conocimiento científico en acceso abierto, sobre los repositorios institucionales como base para el caso de estudio del repositorio Gredos de la Universidad de Salamanca, sobre la literatura gris científica y las tesis doctorales, y sobre las métricas alternativas para medir su uso y su impacto teniendo en cuenta los estudios realizados sobre el tema.

La parte práctica se basó en el análisis cuantitativo de visibilidad, uso y citación de las tesis doctorales de la Universidad de Salamanca leídas en el período comprendido entre 2006 y 2011.

Para obtener la muestra analizada se tuvieron en cuenta, en primer lugar, los datos proporcionados por la base de datos TESEO y los datos proporcionados por Blázquez Ochando (2015) sobre TESEO; para las tesis en acceso abierto se han empleado los datos proporcionados por el propio repositorio.

Para realizar el análisis de uso y citación de las tesis doctorales se utilizaron los datos sobre visitas y descargas extraídos del módulo de estadísticas de la universidad de Tasmania instalado en el repositorio Gredos y los datos sobre citas proporcionados por *Google Scholar Citation*, previamente también se rastrearon y extrajeron los datos sobre las tesis citadas de la base de datos Web of Science.

Se elaboró una base de datos con la muestra de tesis a analizar, en la que se recogió una serie de variables e indicadores sobre el uso, la visibilidad y la citación de las mismas. Para trabajar con los datos se utilizaron Excel y SPSS.

Se realizó un análisis descriptivo y comparativo de las variables objeto de estudio. Las pruebas de distribución de normalidad se realizaron con el test de Kolmogórov-Smirnov. Mediante el coeficiente de correlación de Spearman se estimó la relación entre las variables cuantitativas del estudio. Para analizar la comparación de los indicadores de visibilidad e impacto según las áreas de conocimiento se utilizó el test de Kruskal-Wallis. Para saber entre qué par de grupos había diferencias se realizaron comparaciones por pares (post-hoc) con el test Dunn.

Resultados

Se demuestra que las tesis depositadas en repositorios de acceso abierto alcanzan un grado de visibilidad y uso que no tienen de ninguna forma las tesis que no están en acceso abierto; también se demuestra que las tesis doctorales difundidas a través de repositorios son citadas aunque la relación entre uso y citas recibidas no es proporcional, hecho que permite deducir la falta de buenas prácticas de citación sobre la literatura gris.

Conclusiones

Los repositorios institucionales son un canal válido para la difusión de la literatura gris científica, particularmente de las tesis doctorales.

Se puede establecer un sistema de análisis bibliométrico viable para la medición de la visibilidad, el uso, la citación y el impacto de las tesis depositadas en repositorios institucionales de acceso abierto.

Abstract

Object

The main objective of this work is to prove the benefits of open access as a model of scientific communication for scientific gray literature, in general, and for electronic doctoral theses, in particular. This research work aims to answer the question on visibility, usability and impact of doctoral theses deposited in open access repositories, through a bibliometric case study of Gredos repository of the University of Salamanca.

Method

This research comprises two parts: a theory and practice. For locating and selecting of relevant literature that support both parts of the research it conducted a systematic review of the literature.

In theory part it carried a detailed study of the state of the scientific knowledge on open access, on institutional repositories as a basis for conducting the case of study of Gredos Repository of the University of Salamanca, on grey literature and doctoral theses, and on altmetrics to measure their use and impact according the impact studies about the subject.

Practice part is based on quantitative analysis of visibility, use and citation of doctoral thesis at the University of Salamanca presented in the period between 2006 and 2011.

To obtain the analyzed sample it took into account, first, the data provided by TESEO database and data provided by Blázquez Ochando (2015) about TESEO; for open access PhD theses it used data provided by the repository itself.

To make the use and citation analysis of PhD theses it used data about visits and downloads which were extracted of statistics module of Tasmania University installed in Gredos repository and the citation data provided by Google Scholar Citation, previously it also screened and extracted PhD theses data of Web of Science database.

It developed a database with the sample of PhD theses to analyze, where it collected a number of variables and indicators on the use, visibility and citation. To work with data it used Excel and SPSS.

It carried a descriptive and comparative analysis of the variables under study. The normal distribution tests were performed with the Kolmogorov-Smirnov test. By Spearman correlation coefficient it estimated the relationship between quantitative variables of the study. To analyze the comparison of visibility and impact indicators according to areas of knowledge he used the Kruskal-Wallis tests. To find between what two groups had differences it carried comparisons (post-hoc) pairs with Dunn test.

Results

It is shown that the doctoral theses deposited in open access repositories reach a degree of visibility and use that not open access doctoral theses are not in any way; also it shows that doctoral theses disseminated through repositories are cited although the relationship between use and citations received is not proportional, a fact that allows us to deduce the lack of good practices on grey literature citation.

Conclusions

Institutional repositories are a valid channel to the dissemination of scientific grey literature, particularly the doctoral theses.

It can establish a viable bibliometric analysis system for measuring visibility, use, citation and impact of theses deposited in institutional open access repositories.



MÓDULO 0. INTRODUCCIÓN

Módulo 0. Introducción

El módulo cero está dedicado a la introducción del trabajo de investigación que se presenta y comprende ocho capítulos estructurados de la siguiente forma:

Capítulo 0.1. Objeto de estudio

Capítulo 0.2. Objetivos

Capítulo 0.3. Motivación / Oportunidad de la investigación

Capítulo 0.4. Preguntas de investigación

Capítulo 0.5. Estudio de la bibliografía

Capítulo 0.6. Estructura

Capítulo 0.7. Marco teórico

Capítulo 0.8. Marco metodológico

Se trata de ofrecer una breve introducción a la tesis que engloba el objeto de estudio (capítulo 1) y los objetivos que se persiguen con el mismo (capítulo 2), la motivación para realizar esta investigación (capítulo 3), las preguntas de investigación (capítulo 4), el estudio de la bibliografía resultante de la revisión sistemática de la literatura (capítulo 5), la estructuración de la tesis (capítulo 6), la introducción general al marco teórico (capítulo 7) y la introducción general a la metodología (capítulo 8) donde se indican el nivel y diseño de la investigación, la población y la muestra, y los métodos empleados.

El objetivo del Módulo 0 es situar la tesis doctoral y ofrecer una panorámica de los contenidos que se desarrollan en la misma. La redacción de este módulo introductorio obedece a la intención de que sirva de ayuda para una mejor y más rápida comprensión de los temas que se han estudiado en esta investigación.

Capítulo 0.1. Objeto de estudio

En la era de la explosión de la información, la información se dispersa y está disponible bajo múltiples formas. Una de esas formas en las que se materializa la información científica es la literatura gris, y un tipo de literatura gris que produce resultados de investigación es la tesis doctoral. Estos resultados científicos, hasta hace muy poco tiempo, estaban limitados solo al investigador y al departamento y universidad donde se había presentado la tesis doctoral. Cuando la tesis doctoral se enviaba en formato impreso, era muy difícil que la información que contenía llegara a ser accesible y se diseminara por todo el mundo. La valiosa información de la tesis pasaba inadvertida y no podía ser utilizada por cualquiera.

Hoy en día todavía es frecuente que el investigador no pueda acceder a la información que necesita por la falta de recursos de información, debido al incremento de los precios de suscripción de las revistas científicas, y a la escalada de los precios de los libros y de las bases de datos. Como consecuencia, los investigadores dependen de la calidad de los recursos que están disponibles libremente a través de Internet, siendo las tesis doctorales electrónicas uno de los principales recursos científicos de calidad que pueden estar accesibles libremente.

El objeto de estudio de este trabajo es la literatura gris científica depositada en repositorios institucionales de acceso abierto. Se estudia la visibilidad, el uso y la citación de la literatura gris científica en abierto, concretamente de las tesis doctorales electrónicas depositadas en los repositorios, a través de un estudio bibliométrico de caso del repositorio Gredos de la Universidad de Salamanca.

Gredos es el repositorio institucional de la Universidad de Salamanca, que ofrece en la actualidad (agosto 2016), desde la dirección <http://gredos.usal.es>, más de 105.000 documentos en acceso abierto. Está organizado en cuatro secciones de acuerdo a la heterogénea naturaleza de su contenido: la “Biblioteca Digital” que contiene fondos patrimoniales, el “Repositorio Científico” que contiene fondo de investigación, el “Repositorio Docente” que contiene recursos didácticos y contenido docente, y el

“Archivo Institucional” que contiene documentos históricos y actuales de carácter administrativo (Casaus & Ferreras-Fernández, 2009; Ferreras-Fernández, 2013a).

En la sección “Repositorio Científico” de Gredos, <http://gredos.usal.es/jspui/handle/10366/4756>, además de otros tipos de documentos científicos, se alojan y difunden en acceso abierto las tesis doctorales presentadas en la Universidad de Salamanca a partir del curso 2008-2009; también se alojan, a petición de los autores, tesis doctorales leídas en la Universidad de Salamanca en años anteriores al 2008. En agosto de 2016, el número total de tesis doctorales depositadas en Gredos ascendía a 2097. En la Universidad de Salamanca, al igual que en cualquier otra institución, las tesis doctorales son el principal tipo de literatura gris científica producida por la institución, y una de las principales fuentes de auténtica y detallada información de un tema particular sobre el que un investigador ha llevado a cabo un amplio estudio.

Por otra parte, hay que destacar que actualmente la investigación científica se desenvuelve en un contexto competitivo en el que los investigadores deberán demostrar que su trabajo tiene impacto en su campo de investigación, con el fin de asegurar la financiación, promoción y permanencia. Para maximizar el impacto los investigadores tienen como prioridad obtener resultados primeramente, y seguidamente publicarlos en las revistas del más alto impacto. Así pues, con el desarrollo de Internet se han presentado grandes oportunidades para el investigador que busca aumentar el impacto, al hacer accesible su trabajo inmediatamente en la Web a través de repositorios de acceso abierto.

La importancia que han alcanzado los repositorios como nueva forma de comunicación científica se demuestra por el aumento exponencial de los mismos desde 2006. Según datos del directorio OpenDOAR, <http://www.opendoar.org/>, el número de repositorios en 2006 era de 128 repositorios, en agosto de 2016 este número asciende a más de 3.196, de los que 2.715 son repositorios institucionales.

Los repositorios institucionales (RIs) ofrecen la difusión más amplia posible de toda la oferta de la producción intelectual digital generada en una institución, incrementando la visibilidad y el prestigio de la misma y demostrando su valor a las instituciones de financiación de la investigación (Ferrerías-Fernández, 2010c).

Además los repositorios aumentan la visibilidad de sus investigadores, ampliando la difusión y el uso de sus trabajos; estimulan la innovación, facilitan un análisis cualitativo del trabajo de sus miembros, apoyan las tareas de enseñanza y aprendizaje, sirven de sistema de registro de ideas y ofrecen un catálogo del capital intelectual de la institución. Finalmente, los repositorios institucionales pueden mejorar la comunicación científica y hacer avanzar la investigación permitiendo a los usuarios localizar y recuperar información relevante de forma más rápida y fácil (García-Peñalvo, 2008b; López Medina, 2006).

Un aspecto fundamental a tener en cuenta en esta investigación es que los repositorios institucionales reconocen y dan cabida a los objetos digitales científicos no incluidos en los canales tradicionales de publicación, como sucede con los tipos de literatura gris científica y, en este caso, con las tesis doctorales.

Por ello, las tesis doctorales electrónicas en acceso abierto que están depositadas en repositorios institucionales presentan muchas ventajas:

- hacen visible una parte muy importante de la investigación realizada en las instituciones,
- facilitan el descubrimiento de los autores y de su investigación, posibilitando y aumentando el uso de la misma,
- ayudan al aumento del número de citas del autor y de la institución,
- el tiempo requerido para la difusión de la información académica es mínimo,
- presentan la posibilidad de incorporar archivos en diferentes formatos en forma electrónica, lo cual no es posible en las tesis impresas,

- y son una solución de preservación a largo plazo para las tesis doctorales.

Capítulo 0.2. Objetivos

El principal objetivo de este trabajo es demostrar que la difusión de las tesis doctorales a través de repositorios institucionales aporta más visibilidad y más uso, y, por tanto, más citación e impacto que las tesis doctorales que no son de libre acceso. En el desarrollo de esta hipótesis debe plantearse además otra pregunta: ¿qué factores internos y externos influyen en la visibilidad y el uso de las tesis depositadas en repositorios institucionales de acceso abierto?

Existen estudios que demuestran el aumento de citación de los trabajos de investigación publicados y puestos en acceso abierto frente a aquellos trabajos publicados y que no están en acceso abierto (Atchison & Bull, 2015; Bernius & Hanauske, 2009; W. Cheng & Ren, 2008; Donovan & Watson, 2011; Donovan, Watson, & Osborne, 2014; Evans & Reimer, 2009; Eysenbach, 2006; Frisch, Nathan, Ahmed, Shidham, & Ahmed, 2014; Gargouri et al., 2010; Gaulé & Maystre, 2011; Hajjem, Harnad, & Gingras, 2005; Harnad & Brody, 2004; Koler-Povh, Juznic, & Turk, 2014; Kullman, 2014; Kurtz et al., 2005; Laakso & Björk, 2013; Lawrence, 2001; Mahesh, 2012; McCabe & Snyder, 2014; Michael Norris, Oppenheim, & Rowland, 2008; Piwowar & Vision, 2013; Riera & Aibar, 2013; Xia, Myers, & Wilhoite, 2011; Xia & Nakanishi, 2012; Zhang, 2006).

La novedad de esta investigación radica en el estudio de la literatura gris científica, un campo mucho más difícil de explorar a la hora de obtener datos de uso y citación que aquel de los trabajos publicados por los canales tradicionales de edición (Ferrerías-Fernández, García-Peñalvo, & Merlo-Vega, 2015).

El objetivo principal implica una serie de sub-objetivos:

1. Definir el estado de la cuestión de los repositorios institucionales de acceso abierto tanto desde un punto de vista estratégico-político como tecnológico.
2. Definir el estado de la cuestión de las tesis doctorales electrónicas en acceso abierto.

3. Demostrar que la interoperabilidad de los repositorios (aspecto técnico) y las políticas de mandato afectan positivamente a la visibilidad, el uso y el impacto de la investigación depositada en los repositorios institucionales y por ende a las tesis electrónicas
4. Realizar un estudio cuantitativo en el repositorio Gredos que tome como elemento de estudio las tesis doctorales de la Universidad de Salamanca.
5. Conocer y aplicar métricas alternativas para la realización del estudio.
6. Establecer un modelo bibliométrico viable para la medición del uso y de la citación de las tesis depositadas en acceso abierto.

Capítulo 0.3. Motivación y Oportunidad de la investigación

El trabajo que se presenta se ha desarrollado dentro del Programa de Doctorado Formación en la Sociedad del Conocimiento de la Universidad de Salamanca, a partir del curso 2014-2015.

El Programa de Doctorado Formación en la Sociedad del Conocimiento comienza su andadura en el curso 2013-2014, <http://doctorado.usal.es/es/doctorado/formaci%C3%B3n-en-la-sociedad-del-conocimiento>, se configura en el seno del Instituto de Ciencias de la Educación (IUCE), <https://iuce.usal.es/>, y se define según el Real Decreto 99/2011, de 28 de enero, por el que se regulan las enseñanzas oficiales de doctorado (Ministerio de Educación, 2011).

Se trata de un programa de carácter interdisciplinar y multicultural (García-Peñalvo, 2013a, 2013b, 2013c, 2015) soportado principalmente por los Grupos de Investigación Reconocidos de la Universidad de Salamanca, tales como GRIAL (<http://grial.usal.es>), GITE (<http://gite.usal.es/>), OCA (<http://campus.usal.es/~oca/>), VISUALMED (<http://visualmed.usal.es/>), Robótica y Sociedad (<http://gro.usal.es/>) y E-LECTRA (<http://electra.usal.es/>), de los cuales GRIAL, GITE y OCA son también Grupos de Excelencia de la Junta de Castilla y León (GR47, GR213 y GR319 respectivamente). Además de este grupo de investigadores de la Universidad de Salamanca, el programa se ve completado con un amplio número de investigadores nacionales e internacionales, que conforma la vanguardia de las líneas de investigación relacionadas con la Formación en la Sociedad del Conocimiento. Las líneas de investigación de este programa de doctorado son:

1. Evaluación Educativa y Orientación.
2. Interacción y eLearning.
3. Investigación-Innovación en Tecnología Educativa.
4. Medios de Comunicación y Educación.
5. Medicina y Educación.
6. Robótica educativa.

7. Ingeniería y Educación.

8. Educación y Sociedad de la Información.

Además este trabajo doctoral y la perspectiva de la investigación han sido posibles en gran medida a la formación y al trabajo desempeñado por la doctoranda como bibliotecaria en la Universidad de Salamanca, desde 2001, y anteriormente en la Universidad de Santiago de Compostela (1990-2001), así como a los años dedicados a la defensa y promoción del acceso abierto y de las políticas de información de interés público (Casaus & Ferreras-Fernández, 2009; Ferreras-Fernández, 2010a, 2010c, 2011a, 2011b, 2013a; Ferreras-Fernández & Casaus-Peña, 2008; Ferreras-Fernández & Merlo-Vega, 2015; Merlo Vega & Ferreras-Fernández, 2012; Merlo-Vega & Ferreras-Fernández, 2010). Durante todo ese tiempo de ejercicio de la profesión bibliotecaria, la doctoranda ha tenido la oportunidad de observar de primera mano el impacto de las tendencias a las que se hacen referencia en esta tesis doctoral.

El trabajo realizado en 2007 desde el Servicio de Bibliotecas, <http://bibliotecas.usal.es>, como coordinadora técnica del equipo de trabajo para la planificación, desarrollo y puesta en marcha del repositorio institucional Gredos de la Universidad de Salamanca fue el inicio de una etapa de trabajo para la promoción y desarrollo del acceso abierto en la Universidad de Salamanca a través del repositorio institucional. Previamente, el acceso abierto al conocimiento había suscitado un gran interés personal por lo que la doctoranda estudió e investigó sobre este tema en todas las áreas relacionadas del mismo, tanto teóricas como técnicas: la filosofía del acceso abierto, las políticas, los metadatos, el *software* libre, la digitalización, la preservación, etc.

Seoane-Pardo (2014) en su trabajo de tesis doctoral ofrece una amplia información sobre el programa estratégico "Universidad de Salamanca Digital" y del servicio "Universidad Virtual" de la Universidad de Salamanca, donde se enmarcó el proyecto de repositorio Gredos.

En el año 2007, se creó el Vicerrectorado de Innovación Tecnológica y, al frente del mismo, el Prof. Dr. D. Francisco José García Peñalvo, uno de los directores de esta tesis doctoral, impulsó una estrategia integral de gestión de la innovación tecnológica que afectaba a todos los ámbitos de actuación de la universidad, ya fueran funcionales (docencia, investigación y "tercera misión", gestión económico financiera y de Recursos Humanos, gestión académica y de la investigación) o transversales (cuadro de mando y calidad, capital intelectual y gestión del conocimiento, acceso a la información y los servicios, y gestión y gobierno TIC). Esta visión estratégica es reflejada y compartida por varias instituciones españolas, y así se pone de manifiesto con la publicación del Libro Blanco de la Universidad Digital 2010 (Laviña Orueta & Mengual Pavón, 2008) en el que el Prof. Dr. García Peñalvo participa también como autor (García-Peñalvo, 2008c). En este contexto se explica la creación de la Universidad Virtual como servicio dentro de la propia institución.

El plan estratégico sobre Tecnologías en la Universidad de Salamanca se estructuró en seis ejes (García-Peñalvo, 2009), de los que interesa destacar el eje 4 que correspondía a la apuesta institucional por el Conocimiento en Abierto: la Universidad de Salamanca se adhirió a la Declaración de Berlín sobre el Conocimiento Abierto en Ciencias y Humanidades el 27 de febrero de 2009, y esto supuso una apuesta por la filosofía *open* en diversas áreas, desde el *software* libre hasta los contenidos en abierto (OCW, <http://ocw.usal.es> y repositorio institucional Gredos <http://gredos.usal.es>), pasando por la Innovación abierta y la Ciencia abierta. En esa apuesta por el conocimiento en abierto destaca la creación de Gredos (Gestión de REpositorio DOcumental de la universidad de Salamanca, <http://gredos.usal.es>), repositorio institucional basado en DSpace (<http://www.dspace.org>) y concebido como una herramienta de procesamiento, almacenamiento y recuperación de colecciones en formato digital producidas o alojadas en la universidad.

Aunque desde 2007 se venía trabajando en el proyecto del repositorio institucional, no fue hasta marzo de 2009 que Gredos se presentó oficialmente, como eje central de su estrategia Universidad Digital bajo la filosofía de Acceso Abierto o filosofía Open, congruente con la Declaración de Berlín sobre Acceso Abierto al Conocimiento en las Ciencias y las Humanidades a la que la Universidad de Salamanca se adhirió en Consejo de Gobierno de 27 de febrero de 2009 (Casaus & Ferreras, 2009; García-Peñalvo, 2008b).

Desde los comienzos en esta área de trabajo hasta ahora no han pasado muchos años, sin embargo volviendo la vista atrás se comprueba con satisfacción la madurez alcanzada por la Universidad de Salamanca, llegando a disponer de un repositorio respetable que ocupa una posición destacada en los *ranking* mundiales de repositorios, <http://repositories.webometrics.info/en/search/Rankings/gredos>, y que en muchos momentos ha sido referente en el panorama español e incluso internacional, a pesar de los tiempos difíciles que se sufren con frecuencia cuando el compromiso institucional de la universidad no es lo suficientemente sólido para avanzar adecuadamente en el desarrollo del acceso abierto.

Además, la necesidad de aprender y acrecentar conocimientos sobre digitalización y acceso abierto con el fin de poner en marcha y gestionar adecuadamente el repositorio Gredos fue, quizás, el motivo principal por los que la doctoranda, realizó el Máster en Sistemas de Información Digital, los cursos de Doctorado de Información y Documentación, y el Doctorado en Formación en la Sociedad del Conocimiento, alentada en todo momento por los directores de esta tesis, los profesores doctores D. José Antonio Merlo Vega y D. Francisco José García Peñalvo, referentes siempre en la defensa y el impulso del acceso abierto en la Universidad de Salamanca.

El tema tratado por la doctoranda en el Trabajo de Fin de Máster fue la preservación digital de repositorios institucionales y el caso de estudio fue también el repositorio Gredos (Ferreras-Fernández, 2010c). Al comenzar el Doctorado, el objetivo de esta investigación no era el de realizar un estudio bibliométrico de la visibilidad y el impacto de los contenidos del repositorio Gredos, sino tratar e

investigar temas técnicos como los formatos de metadatos, la interoperabilidad, la preservación y la digitalización en los repositorios de acceso abierto. Con el paso del tiempo y el desarrollo en la investigación, y sin dejar nunca de lado los temas antes referenciados, fueron surgiendo nuevos centros de interés como los estudios bibliométricos sobre uso, visibilidad e impacto de las publicaciones en acceso abierto así como la aplicación al acceso abierto de las nuevas métricas alternativas.

En los talleres y cursos impartidos a los usuarios sobre el repositorio Gredos, siempre surgían dudas y preguntas sobre los beneficios y riesgos de autoarchivar sus trabajos y difundirlos abiertamente a través de Internet. Este fue el impulso para trabajar los temas de la visibilidad y el impacto como temas centrales de esta tesis doctoral. Además como originalidad el estudio se realizaría sobre la literatura gris científica, concretamente sobre las tesis doctorales depositadas en Gredos, ya que la mayoría de los estudios disponibles hasta la fecha se habían realizado casi exclusivamente sobre el impacto de las publicaciones del tipo de artículos en acceso abierto (Hajjem et al., 2005; Harnad & Brody, 2004), y muy pocos habían tratado el tema del impacto de la literatura gris científica en acceso abierto (Friend, 2011; Vincent Larivière, Zuccala, & Archambault, 2008; Prost & Schöpfel, 2014; Schöpfel, Chaudiron, et al., 2014; Schöpfel, Prost, et al., 2014a, 2014b; Schöpfel & Soukouya, 2013; Schöpfel, Stock, Farace, & Frantzen, 2005; Stone & Lowe, 2014).

A través de estos trabajos se ha demostrado que la mayor probabilidad que tienen los artículos en acceso abierto de ser más utilizados y citados en otras investigaciones se debe al hecho de estar en acceso abierto y no se trata de un sesgo de los investigadores de poner en abierto la investigación que ellos crean que tiene más posibilidades de ser utilizada y citada. También los trabajos realizados sobre las tesis doctorales en acceso abierto reconocen los beneficios de incremento de visibilidad, uso e impacto.

Por otro lado, otro de los motivos para seleccionar las tesis doctorales como caso de estudio ha sido que son la única documentación de investigación sobre la que la Universidad de Salamanca estableció un mandato para su depósito y difusión en abierto a través del repositorio Gredos.

Asimismo, se ha considerado que el estudio de la literatura gris científica en abierto constituye un reto por la escasez de estudios realizados sobre este tema. Los tiempos están cambiando y los repositorios son una nueva forma de comunicación científica, por tanto pueden constituir una solución para la publicación (en el sentido de difusión) de la literatura gris no publicada por los medios tradicionales de edición.

El estudio bibliométrico que se ha llevado a cabo ha sido además una oportunidad para estudiar, investigar y aplicar alguna de las nuevas herramientas de medición (métricas alternativas) que han ido surgiendo a raíz del desarrollo e incremento del acceso abierto y de la web social científica.

El hecho de que los beneficios del acceso abierto puedan demostrarse con datos podría permitir a los responsables políticos establecer mandatos para que los autores pongan sus trabajos accesibles libremente en la web, el resultado de los cuales desembocaría en servicios de acceso abierto que permitan una mayor competencia e innovación, que es en definitiva un beneficio para los autores y las instituciones. Demostrar los beneficios del acceso abierto requiere un análisis cuantitativo del efecto que el acceso abierto tiene en la literatura de investigación, en este caso en las tesis doctorales. Para monitorizar el efecto que el acceso abierto ha tenido sobre la comunidad "Tesis doctorales" del repositorio Gredos se han utilizado estadísticas del repositorio y sobre todo *Google Scholar*, <https://scholar.google.es/>, como herramienta alométrica para medir el impacto.

Las tesis doctorales electrónicas del repositorio se benefician ya de los enlaces a la referencia de las mismas y de las métricas de citación que son proporcionadas por los recolectores de los metadatos del repositorio institucional, y que podrían ser también incluidas en cada uno de los ítems y en las listas de referencia de la universidad.

Por último, el estudio bibliométrico sobre uso e impacto que se realiza a lo largo de esta tesis podría ser un modelo válido y replicable en futuros estudios sobre acceso abierto.

En cuanto a la investigación de esta tesis doctoral y los resultados presentados son producto de la aplicación de estudios que comparten muchos de los miembros del GRupo de Investigación en InterAcción y eLearning (GRIAL, <http://grial.usal.es>) de la Universidad de Salamanca. Por ello, se presenta el marco donde se ha gestado y desarrollado esta tesis doctoral durante los últimos tres años.

GRIAL es un Grupo de Investigación Reconocido por la Universidad de Salamanca y Grupo de Excelencia de la Junta de Castilla y León, compuesto por un amplio número de investigadores de diferentes áreas del conocimiento, predominantemente con perfiles técnicos y pedagógicos, además de contar con expertos en gestión de proyectos de eLearning procedentes del campo de las Humanidades, Ciencias Sociales y Ciencias Experimentales, entre otras (García Peñalvo et al., 2012).

GRIAL es también un grupo interdisciplinar, compuesto por miembros permanentes pertenecientes a diferentes áreas del conocimiento, como la ingeniería del *software*, la informática, la educación, las ciencias de la información y la biblioteconomía y, las ciencias sociales, entre otras.

Las principales líneas de investigación desarrolladas por el grupo son:

- Sistemas de aprendizaje interactivos. En esta línea de investigación se estudian temas relacionados con los paradigmas de interacción, las interfaces de usuario, los sistemas colaborativos, la usabilidad, la visualización y el análisis visual.
- Tecnologías para el aprendizaje. Se investiga sobre temas relacionados con plataformas LMS/LCMS, ecosistemas *eLearning*, LO, *software* educativo, estándares y especificaciones *eLearning*, Learning Analytics, MOOC, *mLearning*, PLE (Alier Forment, Casan Guerrero, Conde González, García-Peñalvo, & Severance, 2010; Miguel Á. Conde et al., 2014; Conde González, García-Peñalvo, Rodríguez-Conde, Alier Forment, & García-

Holgado, 2014; Fidalgo-Blanco, Sein-Echaluce Lacleta, Borrás Gené, & García-Peñalvo, 2014; García-Holgado & García-Peñalvo, 2013a, 2013b, 2016; Muñoz et al., 2012)

- Ingeniería del *software*, ingeniería de la Web y arquitectura del software. Se investiga sobre temas relacionados con la Web semántica, las ontologías, la calidad del *software*, los sistemas de gestión del conocimiento, MDA, y los servicios Web (Casany, Alier Forment, Conde, & García-Peñalvo, 2009; M. Á. Conde, García-Peñalvo, Casany, & Alier Forment, 2009; F. J. García-Peñalvo, Conde, Alier Forment, & Casany, 2011; Morales, Gil, & García-Peñalvo, 2007)
- Metodologías eLearning. Esta línea de investigación incluye el estudio del tutorial *Online*, las comunidades de aprendizaje y la comunidad de gestión, y CSCL (García-Peñalvo, 2008d; F. J. García-Peñalvo & Seoane-Pardo, 2015; Seoane-Pardo, 2014; Seoane-Pardo & García-Peñalvo, 2014a, 2014b).
- Calidad y evaluación en educación. Se engloba los programas de evaluación educativa, la evaluación y procesos de enseñanza en entornos virtuales, la guía educativa y vocacional, y las medidas educativas).
- Teoría de la comunicación. Se tratan los tópicos sobre la comunicación interpersonal, la retórica, la interacción lingüística, la comunicación didáctica, y la teoría del lenguaje.
- Gestión estratégica del conocimiento y la tecnología (García-Peñalvo, 2011; García-Peñalvo, García de Figuerola, & Merlo, 2010a; García-Peñalvo, García de Figuerola, & Merlo-Vega, 2010b; García-Peñalvo & Rivera, 2009).
- Humanidades digitales (García-Peñalvo, 2016a; Therón & Fontanillo, 2015; Therón, Seguí, Cruz, & Vaquero, 2014; Therón & Wandl-Vogt, 2014).
- TIC e innovación educativa (García-Peñalvo, 2008a; Laviña Orueta, J., & Mengual Pavón, 2008).
- Ciencias de la Información. En esta línea de investigación se incluyen los temas más directamente relacionados con la investigación de esta tesis doctoral, tales como repositorios,

conocimiento abierto, y bibliotecas digitales (Ferrerías-Fernández, Merlo-Vega, & García-Peñalvo, 2013a, 2013b, 2015; García-Peñalvo, Merlo-Vega, Ferreras-Fernández, et al., 2010; Morales Morgado, Campos Ortuño, Yang, & Ferreras-Fernández, 2014).

Además, las tesis doctorales y los trabajos publicados durante el desarrollo de las mismas en algunas líneas de investigación del Grupo GRIAL han servido de marco de referencia de este trabajo de investigación que se presenta, como son las siguientes tesis:

Seoane-Pardo, A. M. (2014). *Formalización de un modelo de formación online basado en el factor humano y la presencia docente mediante un lenguaje de patrón*. (Tesis Doctoral), Universidad de Salamanca, Salamanca..

Conde, M. Á. (2012). *Personalización del aprendizaje: Framework de servicios para la integración de aplicaciones online en los sistemas de gestión del aprendizaje*. (Tesis Doctoral), Universidad de Salamanca, Salamanca.

González Fernández-Villavicencio, N. (2014). *La rentabilidad de la biblioteca en la web social*. (Tesis Doctoral), Universidad de Salamanca, Salamanca.

Capítulo 0.4. Preguntas de investigación

A partir de los objetivos planteados, a continuación se exponen las preguntas de investigación a las que se quiere dar respuesta en esta tesis doctoral.

1. ¿Las tesis doctorales en acceso abierto tienen más visibilidad y usabilidad que las tesis doctorales que no son de acceso abierto?
2. ¿Las tesis doctorales en acceso abierto son más citadas que las tesis que no son de acceso abierto?
3. ¿Qué factores intrínsecos y extrínsecos influyen en la visibilidad, usabilidad e impacto de las tesis doctorales depositadas en repositorios institucionales de acceso abierto?
4. ¿Cómo se mide la visibilidad, usabilidad e impacto del acceso abierto?
5. ¿Cómo se puede realizar un análisis comparativo de impacto de las tesis en acceso abierto con las tesis que no están en acceso abierto?

Capítulo 0.5. Estudio de la bibliografía

La bibliografía utilizada en esta investigación ha sido el resultado de la una revisión sistemática de la literatura que constituye parte de la metodología de esta tesis doctoral, y uno de los capítulos imprescindibles y más extensos de la misma, recogido en el capítulo 1 del *Módulo I. Estado de la cuestión*. A lo largo de las secciones que componen el capítulo se ha planteado el tópico de las revisiones de literatura en general y de la revisión sistemática, en particular; se han expuesto los objetivos y preguntas de investigación para hallar la literatura pertinente; se ha realizado una breve panorámica sobre la revisión de la literatura, los tipos de revisiones y la revisión sistemática en Ciencias Sociales.

La aplicación del proceso de revisión sistemática llevado a cabo se fundamenta paso a paso con teoría. Los resultados obtenidos y expresados en la sección correspondiente de esta revisión se presentan además acompañados de tablas que contienen todos los datos y resultados del proceso en el Anexo VI. 4.

Se ha realizado una revisión sistemática de la literatura sobre los temas de investigación de la tesis, de los cuales no se hallaron revisiones sistemáticas que se pudieran replicar. El propósito ha sido obtener una bibliografía basada en la mejor investigación disponible en la que fundamentar este trabajo de investigación, tanto para la elaboración del estado de la cuestión como de la metodología y de la práctica.

A través de la búsqueda de la literatura se han recuperado cientos de referencias. Estas han debido de ser filtradas para identificar las que eran útiles para su posterior revisión.

Los primeros resultados se cribaron fundamentalmente de acuerdo a dos criterios de inclusión:

1. Se utilizó la información bibliográfica disponible en las bases de datos para filtrar los resultados por relevancia;

2. Los trabajos a texto completo se obtuvieron por los títulos que parecían más relevantes.

Los detalles de cada trabajo se extrajeron y se evaluaron críticamente por la doctoranda y los directores de esta tesis doctoral para garantizar la coherencia. Los trabajos incluidos se evaluaron críticamente utilizando el marco propuesto en este estudio.

En la Figura 3, *Diagrama de flujo de la selección de estudios para la inclusión en la revisión sistemática* que aparece en el siguiente capítulo se detalla el proceso seguido y las fuentes utilizadas para la obtención de la literatura.

A través de los resultados de la revisión sistemática de la literatura comprobamos que la bibliografía sobre el tema del acceso abierto en general es muy abundante, así como los estudios sobre visibilidad e impacto de la investigación en acceso abierto. En cambio la literatura sobre la relación entre literatura gris científica, acceso abierto e impacto es bastante escasa.

Se hallaron seis revisiones de la literatura sobre algunos de los temas propuestos en las preguntas de investigación sobre acceso abierto que sirvieron para identificar bibliografía sobre las cuestiones planteadas. Además como resultado de la revisión del ámbito de los temas genéricos relacionados con el acceso abierto, repositorios y preservación digital se localizaron 10 Bibliografías de Charles W. Bailey, Jr., muy útiles para la localización de bibliografía en estos temas.

En total, se identificaron 2.192 referencias de las bases de datos WOS, SCOPUS, LISTA y LISA, que se redujeron posteriormente a 552 referencias después de un examen detallado de los títulos y los resúmenes y de la eliminación de duplicados.

Solo 257 referencias cumplieron con los criterios finales para su inclusión en la revisión. El 91% de las referencias son artículos de revistas (233 de 257) procedentes principalmente de 28 revistas pertenecientes al ámbito de la información y la documentación.

Las conferencias ascienden a 17, los capítulos de libros a 3 y a 4 los trabajos clasificados como otro tipo de documentación (carta a editor, noticia de blog, etc.).

Las publicaciones están repartidas en un período de tiempo que abarca 1998 a 2015.

Además, 104 referencias fueron identificadas por métodos complementarios (51 referencias obtenidas de búsquedas manuales y 53 referencias obtenidas a partir de referencias y citaciones de trabajos incluidos). Para *Google Scholar* y los recursos de acceso abierto, el proceso de revisión fue similar aunque adecuando la estrategia de búsqueda a cada uno de los recursos, dando lugar a 883 referencias de las que se seleccionaron 182.

En el capítulo correspondiente a la revisión sistemática de la literatura se explica con detalle el proceso de revisión seguido y se comentan los resultados obtenidos.

Capítulo 0.6. Estructura

Para estructurar este trabajo se han seguido las recomendaciones de Codina (2015) sobre la estructura canónica de una tesis (Tabla 1).

Tabla 1

Estructura de la tesis

MÓDULO	COMPONENTE	CONTENIDO
0	Introducción	Objeto de estudio, objetivos, marco teórico, marco metodológico. Preguntas de investigación. Fuentes
I	Estado de la cuestión	Características más significativas del ámbito de estudio. Principales conceptos y tendencias. Modelos y teorías
II	Sistema de análisis y caso de estudio	El sistema de análisis se refiere a Componentes, funciones y características del protocolo y del sistema de análisis utilizado para la obtención de datos e informaciones. El caso de estudio se refiere a la presentación, descripción y características del ámbito, fenómeno o entidad estudiados.
III	Resultados y discusión	Presentación y discusión de los resultados obtenidos
IV	Conclusiones	Conclusiones separadas y numeradas, relacionadas con las preguntas de investigación
V	Bibliografía	Fuentes
VI	Anexos	Documentación complementaria

Nota: Adaptado de (Codina, 2015). Componentes y contenidos característicos de un trabajo académico.

De acuerdo a este modelo, se ha estructurado esta tesis en seis módulos, considerando cada módulo como un conjunto unitario de piezas que se repiten en la construcción de esta tesis con el fin de hacerla más fácil y manejable a la hora de su lectura y comprensión. Las piezas que componen los módulos se han denominado capítulos, que a su vez se pueden subdividir.

Más adelante, en la introducción correspondiente a cada módulo se ofrece una descripción detallada de la estructura y del contenido de cada uno de ellos.

En este capítulo sobre la estructura de la tesis se expone de forma general y breve cómo se ha organizado la misma.

El módulo 0 está dedicado a la introducción de la tesis. A través de los capítulos en los que este módulo se estructura se definen el objeto de estudio, los objetivos, la introducción al marco teórico y al marco metodológico, se plantean las preguntas de investigación y se realiza un breve estudio de la bibliografía empleada.

El módulo I constituye el Estado del arte o Estado de la cuestión, en el que se presentan las características más significativas de los temas del ámbito de estudio, los principales conceptos y tendencias, y los modelos y las teorías que han servido de base para realizar esta investigación doctoral. El módulo I se compone de 7 capítulos que a su vez se subdividen en epígrafes que especifican el tema tratado en el capítulo.

En el módulo II se presenta el sistema de análisis referido a los componentes, las funciones y las características del protocolo y del sistema de análisis utilizados para la obtención de datos e informaciones. También se trata el caso de estudio que se refiere a la presentación, descripción y características del ámbito estudiado.

En el módulo III se presentan los resultados y discusión del análisis realizado.

En el módulo IV se establecen las conclusiones separadas y numeradas, relacionadas con las preguntas de investigación. También se relaciona la investigación llevada a cabo por la doctoranda, en forma de publicaciones, conferencias y cursos impartidos, durante el Programa de Doctorado Formación en la Sociedad del Conocimiento.

El módulo V está dedicado a la bibliografía utilizada y citada a lo largo de esta tesis doctoral; y el módulo VI, Anexos, contiene toda la documentación complementaria utilizada y referenciada a lo largo de la tesis.

La estructura de la tesis se detalla a continuación y gráficamente en la Figura 1.

Módulo 0. Introducción

Capítulo 0.1. Objeto de estudio

Capítulo 0.2. Objetivos

Capítulo 0.3. Motivación / Oportunidad de la investigación

Capítulo 0.4. Preguntas de investigación

Capítulo 0.5. Estudio de la bibliografía

Capítulo 0.6. Estructura

Capítulo 0.7. Marco teórico

Capítulo 0.8. Marco metodológico

Módulo I. Estado de la cuestión

Capítulo I.1. Revisión sistemática de la literatura

I.1.1. Introducción

I.1.2. Planteamiento del problema

I.1.3. La revisión de la literatura

I.1.4. El proceso de revisión

Capítulo I.2. El conocimiento científico abierto

I.2.1. La comunicación científica

I.2.2. El acceso abierto como nuevo modelo de comunicación científica

I.2.3. Las ventajas del acceso abierto al conocimiento

I.2.4. Contenidos del acceso abierto

I.2.5. Políticas de acceso abierto

I.2.6. Las diez claves del acceso abierto

Capítulo I.3. Repositorios Institucionales (RIs) de acceso abierto

I.3.1. Los repositorios institucionales. Definición y datos

I.3.2. Uso del repositorio y captación del contenido

I.3.3. Factores de éxito de los Repositorios Institucionales

I.3.4. La interoperabilidad y visibilidad de los repositorios

I.3.5. Difusión, visibilidad, uso e impacto de los contenidos de los repositorios

Capítulo I.4. Ecosistemas tecnológicos. El repositorio Gredos

I.4.1. Ecosistemas tecnológicos

I.4.2. El repositorio Gredos

Capítulo I.5. La literatura gris científica. La tesis doctoral

I.5.1. Literatura gris. Definición y datos

I.5.2. Desarrollo de políticas para la gestión de Literatura Gris

I.5.3. La evaluación de Literatura Gris

I.5.4. La tesis doctoral. Las tesis electrónicas en acceso abierto

I.5.5. Beneficios y riesgos de las tesis en acceso abierto: percepción de los autores

I.5.6. El aumento de las e-tesis en acceso abierto. El valor de las e-tesis en los Repositorios

Institucionales

I.5.7. Diseminación de las e-tesis. Redes, recolectores y portales de e-tesis

I.5.8. Acceso a las e-tesis: mejoras para el acceso

Capítulo I.6. Estudios de impacto de la investigación en acceso abierto. Visibilidad, uso y citación de las tesis doctorales

I.6.1. Estudios de impacto de la investigación en acceso abierto

I.6.2. Visibilidad, uso y citación de las tesis doctorales

Capítulo I.7. Bibliometría para el acceso abierto. Métricas alternativas

I.7.1. Bibliometría. Definición y terminología

I.7.2. La comunicación científica en línea y la bibliometría tradicional

I.7.3. Nuevas herramientas y métricas de la ciencia en un escenario de acceso abierto

Módulo II. Sistema de análisis y caso de estudio

Capítulo II.1. Caso de estudio: Tesis doctorales del repositorio Gredos

II.1.1. Colección de Tesis doctorales en el repositorio Gredos

II.1.2. Normativa sobre el depósito de tesis en Gredos

II.1.3. Depósito de las tesis doctorales en Gredos

II.1.4. Diseminación de las tesis doctorales de Gredos: la interoperabilidad

Capítulo II.2. Material y métodos

II.2.1. Muestra

II.2.2. Selección del material

II.2.3. Definición de las variables

II.2.4. Análisis y métodos estadísticos

II.2.4.1. Análisis de las citas en WOS

II.2.4.2. Análisis de visibilidad, uso y citación de las tesis de Gredos

Módulo III. Resultados y discusión

III.1.1. Resultados del análisis de citación en WOS

III.1.2. Resultados del análisis de visibilidad, uso y citación de las tesis de Gredos

Capítulo III.2. Discusión

III.2.1. Cuestiones previas

III. 2.2. WOS como herramienta de búsqueda de tesis citadas

III. 2.3. Medición de la visibilidad, el uso y el impacto de las tesis doctorales en acceso abierto

III.2.4. Google Scholar como herramienta alométrica para la identificación de citas

Módulo IV. Conclusiones

Módulo V. Bibliografía

Módulo VI. Anexos

Anexo 1. Índice de tablas

Anexo 2. Índice de figuras

Anexo 3. Revisión sistemática (carpeta en Tesis versión final)

Anexo 4. Base de datos de las tesis doctorales analizadas

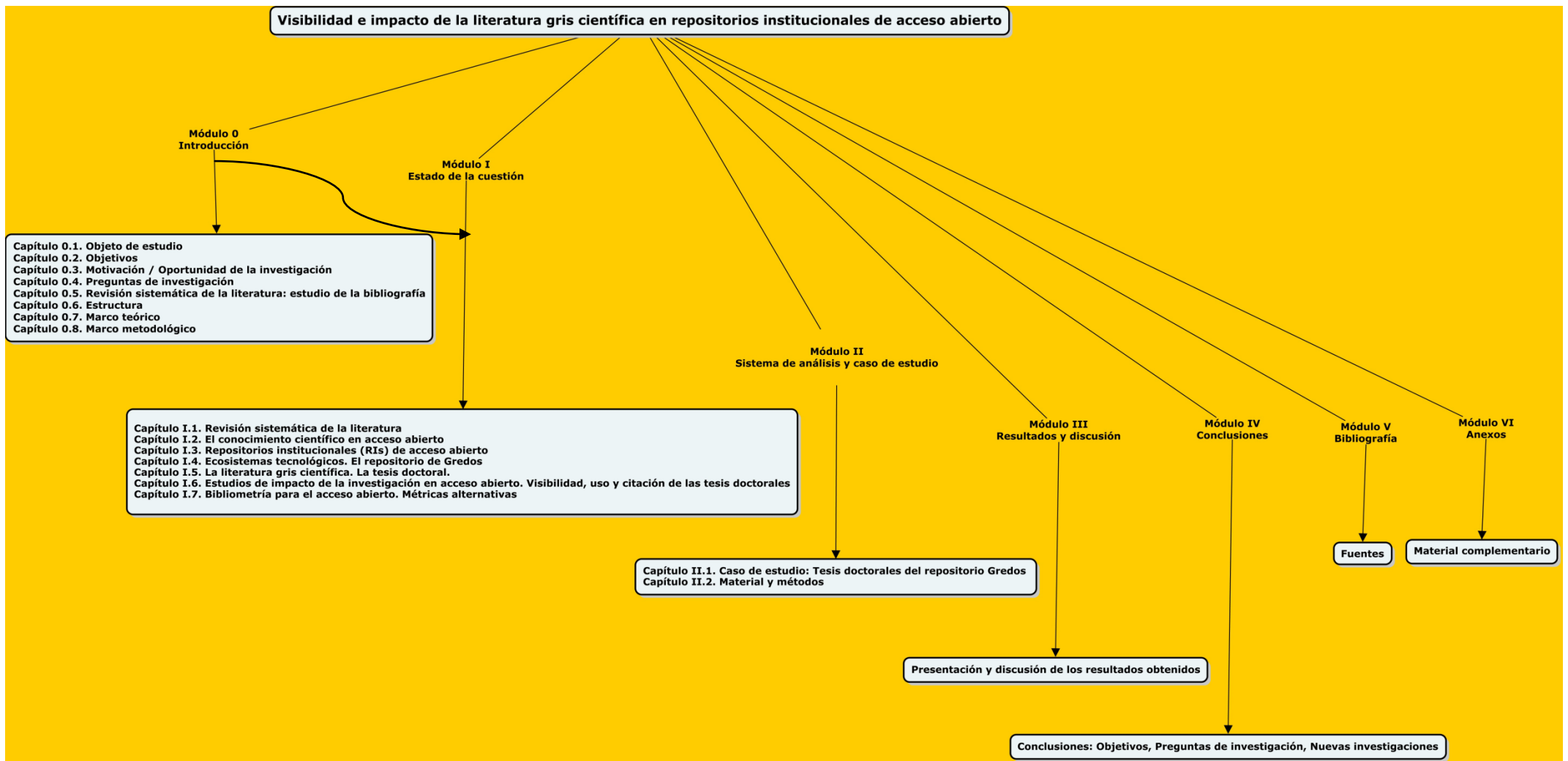


Figura 1. Estructura de la tesis. Fuente: elaboración propia realizada con *Cmap Tools*

Capítulo 0.7. Marco teórico

La actual era de la información ha cambiado el concepto tradicional de biblioteca. Hoy en día se han popularizado los conceptos de biblioteca digital, y de biblioteca electrónica y en cierta medida también el concepto de repositorio, tanto institucional como temático (Barrueco Cruz & García Testal, 2009; Calzada Prado, 2010; Keefer, 2007; Kennan & Wilson, 2006; Lynch, 2003; Remedios Melero, 2008; Nicholas, Rowlands, Watkinson, Brown, & Jamali, 2012; Rovira, Marcos, & Codina, 2007).

La aplicación y desarrollo de la información y la comunicación (TIC) trajo consigo una revolución en las colecciones y en los servicios prestados por las bibliotecas. El conocimiento que tradicionalmente se mantenía cerrado en un libro, revista, actas de congresos, informes de investigación, tesis doctorales, etc. ahora tiene la oportunidad de romper las barreras y llegar al máximo de usuarios en un tiempo mínimo. El acceso abierto ha acelerado el proceso de difusión de la información para cualquier persona, en cualquier momento y en cualquier lugar.

La forma de comunicar la ciencia ha sufrido un cambio radical en los últimos años. Para dar visibilidad a la investigación, así como soslayar las grandes sumas de dinero que numerosas instituciones emplean en la adquisición, sobre todo, de publicaciones periódicas científicas especializadas (*journals*) en soporte papel o electrónico, surgen los Depósitos Digitales en acceso abierto, fruto del movimiento *Open Archives Initiative* (OAI). Estos constituyen una herramienta alternativa para difundir conocimiento, pues han sido concebidos para compartir información, paliando las desmesuradas inversiones, que en ocasiones -y pese al establecimiento de redes, consorcios y alianzas cooperativas- no se pueden realizar por falta de medios, así como aportación para frenar la brecha digital, difundiendo gratuitamente la ciencia (Borgoños Martínez, 2010).

El movimiento *Open Access* trajo consigo el desarrollo de Repositorios Institucionales que emergieron como una nueva estrategia en 2002. La explosión de la información, la falta de presupuestos de las bibliotecas, los problemas de espacio, la alta demanda de información, el aumento de la cantidad de suscripciones a revistas, obligaron a las bibliotecas a buscar diferentes maneras para recoger, almacenar y difundir la información entre los usuarios. Para resolver estos problemas, el concepto de repositorios institucionales se desarrolló entre las instituciones académicas de todo el mundo. Los repositorios institucionales contienen activos digitalizados de la institución como libros, revistas, actas de congresos, tesis y disertaciones, informes de investigación, conjuntos de datos, recortes de periódicos, bases de datos, etc. El movimiento de acceso abierto que comenzó con la Iniciativa de Acceso Abierto de Budapest (BOAI) en el año 2002 ha extendido sus alas a casi todos los países desarrollados y en vías de desarrollo y con diversas organizaciones internacionales que enmarcan las directrices de la política para el desarrollo del acceso abierto (BOAI, 2002).

Las tecnologías de la información y especialmente Internet han cambiado el modo en el que se produce, publica y comunica la información hacia modelos en los que la información se genera fundamentalmente en formato digital y se consume a través de medios en línea. Gracias a Internet, la distribución de documentos electrónicos es fácil, simple y barata. Una de las formas más importantes y eficientes para el intercambio de conocimientos es la filosofía del acceso abierto al conocimiento, porque el conocimiento es absolutamente indispensable para la evolución social en un mundo globalizado (García-Peñalvo, García de Figuerola, & Merlo, 2010).

Las tecnologías de la información y la digitalización no son un fin en sí mismas, sino que son simplemente herramientas. Sin embargo, son herramientas tan potentes que están cambiando el modo de hacer las cosas y que por tanto no pueden ser ignoradas. Un uso adecuado de las mismas permitirá a las organizaciones ocupar una posición privilegiada. Se trata de un proceso continuo de transformación de la propia organización. El paso al mundo digital demanda una

reingeniería de todos los procesos e incluso un replanteamiento de los objetivos (Llorens Largo, 2014).

Según Peter Suber (2004) el movimiento de acceso abierto depende de las tecnologías de Internet y del consentimiento del autor o titular de los derechos de autor (Arlitsch and O'Brien, 2012). Los repositorios institucionales fueron uno de los productos de este movimiento; estos capturan los resultados intelectuales del profesorado, de los investigadores y de los alumnos de las instituciones académicas y aseguran un acceso perpetuo y libre a estos resultados, salvando los periodos de embargo u otras restricciones del editor.

Así pues, los repositorios institucionales (RIs) a menudo, y prácticamente en todos los casos, difunden tesis electrónicas inéditas (Schöpfel, 2013) porque las tesis doctorales no dejan de ser el trabajo de investigación primigenio de todo investigador. Hasta ahora, esta falta de difusión de estos trabajos de investigación inéditos hizo que fueran considerados como literatura gris. En este sentido, los repositorios son una poderosa herramienta para la difusión de la investigación llevada a cabo en las universidades a través de la difusión de las tesis defendidas en el seno de la institución, y por ello se considera que las tesis en abierto deberían dejar de calificarse como literatura gris si se difunden en acceso abierto.

Durante la última década los RIs están alcanzando el peso suficiente para ser fuentes viables para las publicaciones de investigación. También parecen contribuir de manera significativa a las tasas de citación del autor. Los repositorios institucionales pueden desempeñar un papel crucial en la medición de resultados de la investigación, y a su vez pueden afectar a los *ranking* universitarios (The Times Higher Education, 2010).

En relación con el tema del aumento de la visibilidad y de la citación de los trabajos en acceso abierto a través de repositorios (vía verde) existe un número importante de estudios de prestigiosos autores comprometidos con el acceso abierto, como Stevan Harnad y Tim Brody

(2004), Xia y Nakanishi (2012), Lawrence (2001), Piwovar y Vision (2013). La mayoría de estas investigaciones han demostrado que los documentos de acceso abierto tienen una ventaja en el número de citas (Gargouri et al., 2010).

En el presente estudio, se plantea otra pregunta: ¿Tienen las tesis doctorales en acceso abierto una ventaja en citación sobre las tesis doctorales que no están en acceso abierto? De acuerdo con Jones y Andrew (2005) la práctica de hacer versiones digitales de tesis doctorales disponibles en línea está creciendo a nivel internacional. Los Repositorios de Tesis y Disertaciones Electrónicas (ETDs) son habituales en las universidades de todo el mundo. Los datos de estos repositorios sugieren un aumento espectacular en el uso y la citación de tesis doctorales en la actividad de la investigación actual, y esto se relaciona con el aumento del descubrimiento de recursos y la facilidad de acceso que confieren los sustitutos digitales.

En la Universidad de Salamanca, el depósito de tesis doctorales en el repositorio fue regulado por mandato en 2008. Desde 2011 el “Real Decreto 99/2011” también incluye en su capítulo 14 la obligación de difundir las tesis en acceso abierto.

El presente estudio se desarrolla en este contexto. Se considera que tanto el crecimiento del movimiento de acceso abierto, como el desarrollo técnico como por ejemplo los nuevos esquemas de metadatos para facilitar la interoperabilidad y la preservación a largo plazo (Ferrerías-Fernández, Merlo-Vega, et al., 2015; García-Peñalvo et al., 2010; Morales Morgado et al., 2014), o como la adopción de políticas de mandato de acceso abierto por parte de las instituciones (Harnad, 2013; Harnad, Carr, & Swan, 2009; Harnad, 2008, 2009; Harnad, 2006; Vincent-Lamarre, Boivin, Gargouri, Lariviere, & Harnad, 2014) aumentan la visibilidad de la investigación depositada en los repositorios institucionales, incrementando de esta forma la citación y el impacto.

En la parte teórica de esta investigación se ha realizado un estudio pormenorizado del estado de la cuestión sobre el conocimiento científico en acceso abierto, sobre los repositorios institucionales, como base para el caso de estudio del repositorio Gredos de la Universidad de Salamanca, sobre la literatura gris científica y sobre el impacto de la investigación en acceso abierto.

Se ha identificado la bibliografía especializada en las siguientes áreas:

- El conocimiento científico en acceso abierto.
- Repositorios institucionales: políticas de depósito en acceso abierto, interoperabilidad de los repositorios y diseminación de los contenidos, visibilidad y *ranking* web de los repositorios.
- Citación e impacto de la investigación en acceso abierto: estadísticas de uso, índices de citas e impacto y métricas alternativas.

Para ello se ha llevado a cabo una revisión sistemática bibliográfica de las principales fuentes de la literatura científica en la web, en los catálogos de bibliotecas y en las bases de datos especializadas. Determinando, dentro de las áreas anteriormente expuestas, palabras clave y autores relevantes en la materia. La gestión de la bibliografía resultante se ha realizado con *Mendeley*, y se ha utilizado *APA* como estilo de citación.

Se han revisado los estudios sobre visibilidad, citación e impacto de la investigación en acceso abierto, a través de bibliografía sobre el tema.

Se ha estudiado la relación entre buenas prácticas de gestión técnica (interoperabilidad) y política (mandato) y el aumento de visibilidad e impacto de los repositorios institucionales de acceso abierto.

Se han estudiado los aspectos técnicos en la implementación y mejoras en el repositorio Gredos.

Capítulo 0.8. Marco metodológico

La investigación que se presenta pretende lograr los objetivos y sub-objetivos expresados en el capítulo 0.2. donde se ha indicado que el objetivo fundamental de esta tesis doctoral consiste en demostrar las ventajas que reporta el acceso abierto en relación al aumento de la visibilidad, el uso y la citación para la literatura gris científica, utilizando un estudio de caso de las tesis doctorales de la Universidad de Salamanca depositadas en el repositorio Gredos.

La presente investigación se compone de dos partes: una teórica y otra práctica.

En la parte teórica se realizó un estudio pormenorizado del estado de la cuestión sobre el conocimiento científico en acceso abierto, sobre los repositorios institucionales como base para el caso de estudio del repositorio Gredos de la Universidad de Salamanca, sobre la literatura gris científica y las tesis doctorales, y sobre las métricas alternativas para medir su uso y su impacto teniendo en cuenta los estudios realizados sobre el tema. Para la localización y selección de la literatura pertinente en la que sustentar las dos partes de la investigación se llevó a cabo una revisión sistemática de la literatura.

La parte práctica de la tesis doctoral se basa en un estudio de caso, en concreto en la colección de tesis doctorales del repositorio Gredos. Un estudio o análisis de caso es un instrumento o método de investigación que implica una investigación procesual, sistemática y profunda del caso seleccionado.

En cuanto a los objetivos que el estudio de caso puede tratar son los siguientes:

- Producir un razonamiento inductivo. A partir del estudio, la observación y recolección de datos establecer hipótesis o teorías.
- Puede producir nuevos conocimientos al lector, o confirmar teorías que ya se sabían.

- Hacer una crónica, un registro de lo que va sucediendo a lo largo del estudio.
- Describir situaciones o hechos concretos.
- Proporcionar ayuda, conocimiento o instrucción a caso estudiado.
- Comprobar o contrastar fenómenos, situaciones o hechos.
- Pretende elaborar hipótesis.
- Es decir, el estudio de caso pretende explorar, describir, explicar, evaluar y/o transformar.

Para la selección de un caso, puede atenderse al carácter representativo de un caso concreto, aunque la intención del estudio de caso no sea precisamente la de generalizar datos (puede ser que la intención sea transformar esa realidad, y no generalizar a otros casos).

En cuanto a la identificación del problema a investigar, para Yin (1993) los interrogantes de investigación o preguntas de estudio son el primer elemento del diseño de cualquier investigación. Estos interrogantes identifican el problema central de la investigación e indican qué metodología de investigación será la más adecuada. Según este autor los interrogantes "cómo" y "por qué" son los más indicados para una metodología de estudio de caso.

El estudio que propone esta investigación doctoral es analizar las tesis doctorales depositadas en el repositorio institucional Gredos, cómo se utilizan (visibilidad, uso y citación) y por qué (factores que influyen en el uso). De manera que permita encaminar la investigación en la dirección correcta mostrando aquello que es necesario observar para obtener evidencias y evitar "recogerlo todo". El estudio se basó en el análisis cuantitativo de visibilidad, uso y citación de las tesis doctorales de la Universidad de Salamanca leídas en el período comprendido entre 2006 y 2011.

De acuerdo con Yin (1993) el análisis de los datos no se produce nunca a partir del vacío. Es necesario un marco teórico donde situar las hipótesis provisionales de solución del problema detectado o, como mínimo, algunas presuposiciones sobre cómo tratarlo. La teoría juega un papel esencial en el desarrollo de las investigaciones de estudio de caso para contrastar una teoría o para desarrollar una nueva (Reigeluth & Frick, 1999).

En cambio, otros autores como Stake (1995) proponen aplicar un diseño de investigación menos estructurado y mucho más inductivo utilizando una estructura conceptual por temas que estará presente en las cuatro fases de un diseño clásico ayudando a concretar los problemas, guiando la recolección de datos y proponiendo soluciones.

Por lo que respecta a la recolección de datos, Yin (1994) establece hasta seis métodos de obtención de datos o "fuentes de evidencias": documentación, documentos de archivo, entrevistas, observación directa, observación participante y objetos físicos. Este autor propone que para el análisis de datos e interpretación de los resultados es necesario especificar previamente el desarrollo de la investigación: cómo se relacionarán los datos obtenidos con las proposiciones o hipótesis definidas y qué criterios serán utilizados para interpretar los resultados.

Englobando estos componentes Yin (1994) propone definir un marco teórico preliminar sobre el problema estudiado a partir de los conocimientos previos disponibles en la bibliografía científica. Es necesario hacer esta definición teórica antes de iniciar la recolección de datos porque ayudará en el diseño de la investigación y en la interpretación de los resultados.

Los pasos para la elaboración de este estudio de caso se han basado en las siguientes premisas:

1. Elegir un fenómeno de estudio y describir de la forma más completa que se pueda.
2. Recolectar la mayor cantidad de información posible respecto al objeto de estudio.
3. Estructurar y organizar la información.

4. Definir o desarrollar el marco teórico.
5. Confrontar los datos recopilados con el marco teórico.
6. Seleccionar la información útil del marco teórico y los datos.
7. Escribir una serie de preguntas que servirán como guía para el estudio de caso.
8. Definir los principales aspectos o temas de la investigación.
9. Analizar la información seleccionada y analizarla de acuerdo con preguntas formuladas y a los aspectos relevantes a investigar.
10. Hacer una breve conclusión del caso y escribir la bibliografía consultada.

El proceso de aplicación de esta metodología a esta tesis doctoral se explica a continuación, en los siguientes párrafos. Para ello, también conviene ubicar en un marco temporal las fases en las que se ha desarrollado la investigación.

Para alcanzar la consecución de estos, se siguió una metodología de trabajo propia de toda investigación científica. Desde que se inscribió la tesis doctoral hasta que se estableció la redacción definitiva de la misma, se ha estudiado, analizado y sintetizado toda la literatura de interés sobre visibilidad, uso, impacto de la literatura gris científica y sobre las tesis doctorales en acceso abierto.

El proyecto de la presente tesis se inscribió en la Universidad de Salamanca el 20 de mayo de 2015, aunque ya se venía trabajando en esta área del acceso abierto desde 2012. En la planificación temporal presentada en el Plan de investigación de la doctoranda (Ferrerías-Fernández, 2015) se indicaba el calendario previsto en la elaboración de la tesis doctoral, que sufrió algunos retrasos de ejecución a lo largo del desarrollo de esta investigación.

El cronograma de esta tesis doctoral ha sido el siguiente:

1) Estudio de la bibliografía. Se llevó a cabo un estudio detallado de la bibliografía existente a través de la revisión sistemática de la literatura científica sobre los temas planteados. Toda la

bibliografía resultante se gestionó con *Mendeley*. El período de tiempo previsto para la realización de esta fase del trabajo era hasta marzo de 2015. Aunque la recopilación de bibliografía de forma manual se había iniciado en 2012, el proceso de revisión sistemática se llevó a cabo desde agosto de 2014 hasta diciembre de 2015, continuando en 2016 la actualización de la misma.

2) La elaboración de la base de datos de la muestra a partir de TESEO y el repositorio Gredos, se había realizado de enero a abril de 2013.

3) La recopilación de datos de los indicadores sobre uso, visibilidad y citación utilizados en la muestra se realizó en agosto de 2014.

4) El estudio de los datos de la muestra: normalización y estudio con Excel. En 2015 se hicieron algunos análisis sobre una muestra aleatoria de tesis doctorales en Gredos, con el fin de presentar el estudio en algún congreso y de publicarlo en alguna revista científica. Posteriormente se analizó la muestra completa.

5) Elaboración de una propuesta de análisis bibliométrico de aplicación a las tesis depositadas en repositorios institucionales de acceso abierto con el fin de medir el uso, la visibilidad, la citación y el impacto de las mismas. Periodo: abril/mayo 2015.

6) La redacción de los capítulos de la tesis cuya fecha de finalización prevista era junio de 2015, se demoró y se finalizó en julio de 2016.

En cuanto a los contenidos, se ha establecido un desarrollo gradual de los mismos para ofrecer en primer lugar las informaciones básicas necesarias para presentar el estudio de caso que se ha analizado en esta investigación. En el capítulo 0.6.sobre la estructura de la tesis se ha detallado la disposición de los contenidos. Se ha pretendido, a medida que se avanza en la lectura de la tesis doctoral establecer un desarrollo gradual de los mismos. Se comienza con un

módulo 0 introductorio, seguido del módulo 1 sobre el estado del arte o de la cuestión donde se van desarrollando los temas a tratar a partir del estudio de la literatura obtenida mediante una revisión sistemática.

Para la organización de esta tesis, se ha elegido el estilo APA (American Psychological Association, 2010), por su amplia utilización en el área de la Ciencias Sociales, y aplicado su manual de publicación tanto en algunos aspectos de forma para la redacción del texto, la elaboración de tablas y figuras, como para la citación y referencia.

En cuanto a la gestión de las referencias bibliográficas, se ha utilizado *Mendeley*, que es una aplicación web y de escritorio, propietaria y gratuita. Permite gestionar y compartir referencias bibliográficas y documentos de investigación, encontrar nuevas referencias y documentos y colaborar en línea. *Mendeley* combina *Mendeley Desktop*, una aplicación de gestión de referencias bibliográficas y de documentos en formato PDF (disponible para *Windows*, *Mac* y *Linux*) con *Mendeley web*. Ambas aplicaciones son la base sobre la cual se ha creado y desarrollado una red social en línea de investigadores, similar a *lastfm* en cuanto a funcionalidades para identificar, capturar, etiquetar, clasificar y referenciar artículos científicos y académicos. Su comunidad está formada por 3 millones de usuarios y dispone de una base de datos con más de 100 millones de referencias. Desde 2013 pertenece al grupo editorial Elsevier.

Para obtener la muestra analizada se tuvieron en cuenta, en primer lugar, los datos proporcionados por la base de datos TESEO, los datos proporcionados por Blázquez Ochando (2015) sobre TESEO, y para las tesis en acceso abierto los datos proporcionados por el propio repositorio.

Para realizar el análisis de uso y citación de las tesis doctorales se utilizaron los datos sobre visitas y descargas extraídos del módulo de estadísticas de la Universidad de Tasmania instalado en el repositorio y los datos sobre citas proporcionados por *Google Scholar Citation*,

también se rastrearon y extrajeron los datos sobre las tesis citadas de la base de datos *Web of Science*.

Se elaboró una base de datos con la muestra de tesis a analizar, en la que se recogió una serie de variables e indicadores sobre el uso, la visibilidad y la citación de las mismas. Para trabajar con los datos se utilizó el programa IBM *SPSS* para *Windows*, Version 23.0 (IBM Corp, Armonk, NY), y la hoja de cálculo *Excel* 2007.

Se realizó un análisis descriptivo y comparativo de las variables objeto de estudio. Las pruebas de distribución de normalidad se realizaron con el test de Kolmogórov-Smirnov. Mediante el coeficiente de correlación de Spearman se estimó la relación entre las variables cuantitativas del estudio. Para analizar la comparación de los indicadores de visibilidad e impacto según las áreas de conocimiento se utilizó el test de Kruskal-Wallis. Para saber entre qué par de grupos había diferencias se realizaron comparaciones por pares (post-hoc) con el test Dunn.

A partir del análisis de los resultados se han podido establecer comparaciones y relaciones relativas a:

- El volumen de investigación doctoral realizada en la Universidad de Salamanca en cada una de las áreas de conocimiento por año.
- Comparación de las tasas de uso entre las distintas áreas de conocimiento a las que pertenecen las tesis.
- Comparación de las tasas de citación entre las distintas áreas de conocimiento.
- Relación entre las variables visitas y descargas que miden el uso.
- Relación entre las variables que miden el uso con la variable citación que mide el impacto.

Asimismo se ha podido establecer un sistema de análisis bibliométrico viable para la medición del uso y la citación de las tesis depositadas en repositorios institucionales de acceso abierto, que puede ser replicado en estudios similares.

Finalmente se han establecido las conclusiones de la investigación.



MÓDULO I. ESTADO DE LA CUESTIÓN

Módulo I. Estado de la cuestión

El módulo 1 presenta las características más significativas de los temas del ámbito de estudio, los principales conceptos y tendencias, y los modelos y las teorías que han servido de base para realizar esta investigación doctoral.

El módulo 1 se compone de 7 capítulos que a su vez se subdividen en epígrafes que especifican el tema tratado en el capítulo.

En el primer capítulo se realiza una panorámica sobre las revisiones de la literatura y se realiza una revisión sistemática de la literatura como parte de la metodología de esta tesis doctoral. La bibliografía obtenida de la misma fundamenta el resto de los capítulos del estado de la cuestión.

El capítulo 2 está dedicado al conocimiento científico abierto. Se trata por una parte de dibujar el marco en el que se desarrolla el movimiento de acceso abierto, sus ventajas, sus contenidos y sus políticas; y por otra parte de constatar que el acceso abierto representa un nuevo modelo de comunicación científica.

El capítulo 3 está dedicado a los repositorios institucionales (RIs) de acceso abierto. Se ofrece una panorámica sobre los mismos. Se presentan como una herramienta idónea y válida para la difusión del conocimiento a través de la definición, de las características y de las ventajas que ofrecen en cuanto a que dotan de visibilidad, usabilidad e impacto a los contenidos en ellos depositados.

El capítulo 4 está dedicado a los ecosistemas tecnológicos considerados como una nueva arquitectura en red y trabajo colaborativo. Se realiza una aproximación al concepto, características y objetivos del ecosistema tecnológico. Se presenta el repositorio Gredos como parte del ecosistema tecnológico de la Universidad de Salamanca dentro del Proyecto Universidad Virtual.

En el capítulo 5 se centra en la literatura gris científica, y específicamente en las tesis doctorales. Se realiza un estudio de situación de la literatura gris (definición, políticas de gestión, y evaluación) y de las tesis doctorales electrónicas en acceso abierto depositadas en los repositorios institucionales, de forma que permita enmarcar adecuadamente el tema de central de este trabajo.

El capítulo 6 se ha dedicado, por un lado, al comentario y análisis de los estudios de impacto de la investigación en acceso abierto, localizados a través de la revisión sistemática de la literatura, y, por otro, al análisis de los estudios sobre visibilidad, uso y citación de las tesis doctorales.

El capítulo 7 viene a cerrar el módulo I tratando en líneas generales el tema de la bibliometría para el acceso abierto y haciendo referencia a las nuevas métricas alternativas.

La estructura del módulo es la siguiente:

Módulo I. Estado de la cuestión

Capítulo I.1. Revisión sistemática de la literatura

I.1.1. Introducción

I.1.2. Planteamiento del problema

I.1.3. La revisión de la literatura

I.1.4. El proceso de revisión

Capítulo I.2. El conocimiento científico abierto

I.2.1. La comunicación científica

I.2.2. El acceso abierto como nuevo modelo de comunicación científica

I.2.3. Las ventajas del acceso abierto al conocimiento

I.2.4. Contenidos del acceso abierto

I.2.5. Políticas de acceso abierto

I.2.6. Las diez claves del acceso abierto

Capítulo I.3. Repositorios Institucionales (RIs) de acceso abierto

- I.3.1. Los repositorios institucionales. Definición y datos
- I.3.2. Uso del repositorio y captación del contenido
- I.3.3. Factores de éxito de los Repositorios Institucionales
- I.3.4. La interoperabilidad y visibilidad de los repositorios
- I.3.5. Difusión, visibilidad, uso e impacto de los contenidos de los repositorios

Capítulo I.4. Ecosistemas tecnológicos. El repositorio Gredos

- I.4.1. Ecosistemas tecnológicos
- I.4.2. El repositorio Gredos

Capítulo I.5. La literatura gris científica. La tesis doctoral

- I.5.1. Literatura gris. Definición y datos
- I.5.2. Desarrollo de políticas para la gestión de Literatura Gris
- I.5.3. La evaluación de Literatura Gris
- I.5.4. La tesis doctoral. Las tesis electrónicas en acceso abierto
- I.5.5. Beneficios y riesgos de las tesis en acceso abierto: percepción de los autores
- I.5.6. El aumento de las e-tesis en acceso abierto. El valor de las e-tesis en los Repositorios Institucionales
- I.5.7. Diseminación de las e-tesis. Redes, recolectores y portales de e-tesis
- I.5.8. Acceso a las e-tesis: mejoras para el acceso

Capítulo I.6. Estudios de impacto de la investigación en acceso abierto. Visibilidad, uso y citación de las tesis doctorales

- I.6.1. Estudios de impacto de la investigación en acceso abierto
- I.6.2. Visibilidad, uso y citación de las tesis doctorales

Capítulo I.7. Bibliometría para el acceso abierto. Métricas alternativas

- I.7.1. Bibliometría. Definición y terminología
- I.7.2. La comunicación científica en línea y la bibliometría tradicional
- I.7.3. Nuevas herramientas y métricas de la ciencia en un escenario de acceso abierto

Capítulo I.1. Revisión sistemática de la literatura

I.1.1. Introducción

En la metodología de esta investigación se ha incluido la revisión sistemática como procedimiento para la recuperación, análisis y evaluación de la literatura pertinente en la que se basa tanto la parte teórica de esta tesis como la parte práctica. Las revisiones sistemáticas son un tipo de investigación científica que tiene como propósito integrar de forma objetiva y sistemática los resultados de los estudios empíricos sobre un determinado problema de investigación, con objeto de determinar el estado de la cuestión en ese campo de estudio.

Lluís Codina (2015) comenta en su blog que los trabajos de revisión tienen una gran aceptación en el ámbito académico anglosajón y últimamente se está utilizando cada vez más el término *revisión sistemática* para denominar a las revisiones bibliográficas que deben preceder o justificar un estado de la cuestión o en general cualquier estudio de revisión. La metodología es una parte muy importante en un trabajo académico y siempre es necesario llevar a cabo los llamados *análisis bibliográficos* para poder emprender cualquier trabajo de una mínima envergadura como lo son las tesis doctorales. Codina hace una precisión terminológica y considera inadecuado el término *análisis bibliográfico* ya que, literalmente, sería el análisis, por ejemplo, del número de veces que una obra es citada por otras, o cualquier otra métrica del mundo de la bibliometría. Lo que casi todo académico hace en realidad cuando lleva a cabo esta actividad es un análisis de la producción científica más relevante en su campo.

Según Randolph (2009) escribir una revisión de la literatura defectuosa es una de las muchas maneras de hacer fracasar una tesis (Randolph, 2009). Si la revisión de la literatura es deficiente, el resto de la tesis también puede verse como imperfecta, porque "un investigador no puede realizar una investigación significativa sin conocer primero la literatura en el campo de estudio" (Boote & Beile, 2005).

Llevar a cabo una revisión de la literatura es una forma de demostrar el conocimiento sobre un campo de estudio particular, incluyendo el vocabulario, las teorías, las variables y fenómenos clave, los métodos y la historia. Al llevar a cabo una revisión de la literatura también se obtiene información sobre los investigadores y grupos de investigación influyentes en el área. Según algunos autores, la revisión de la literatura puede ser, finalmente y con algunas modificaciones, un "documento académico legítimo y publicable" (Campbell & Menk, 2003).

A lo largo de las secciones de este capítulo se hace el planteamiento del problema sobre las revisiones de la literatura, en general y de la revisión sistemática, en particular. Se plantean los objetivos perseguidos y las preguntas de investigación a resolver. Como preámbulo al estudio se realiza un breve estado de la cuestión sobre la revisión de la literatura, los tipos de revisiones y la situación de la revisión sistemática en Ciencias Sociales, Humanidades y Bibliotecas y Ciencias de la Información y Documentación. La aplicación del proceso de revisión sistemática llevado a cabo se fundamenta paso a paso con teoría. Los resultados obtenidos y expresados en la sección correspondiente de esta revisión se presentan además acompañados de tablas que contienen todos los datos y resultados del proceso en el Anexo VI.4.

I.1.2. Planteamiento del problema

Cualquier investigación primaria debe estar precedida de una revisión sistemática (Xu, Kang, & Song, 2015). Sin embargo existen lagunas de conocimiento respecto al uso de este método en lo que respecta a la literatura en las áreas de la biblioteconomía y de las ciencias de la información, y en general en el área de las ciencias sociales. Al igual que estos autores también se plantea la necesidad de un mayor conocimiento acerca de cómo se utiliza este método en estas áreas, por ejemplo, en información bibliográfica, en vacíos temáticos, y en evaluación de la calidad. Dentro de esos vacíos temáticos se ha detectado la falta de revisiones sistemáticas en los temas de investigación tratados en esta tesis doctoral: el acceso abierto al conocimiento, la visibilidad, la

citación y el impacto, y de todos los temas directamente relacionados con el acceso abierto como son los repositorios, la interoperabilidad de los mismos, y las políticas a favor del acceso abierto. Tampoco existen revisiones sistemáticas sobre literatura gris científica, en concreto sobre tesis doctorales y la difusión y la citación de las mismas. Es verdad que existen informes, directrices, numerosas listas de bibliografía sobre todos estos temas pero ninguna revisión sistemática de la literatura que podamos replicar en nuestra investigación.

Xu, Kang y Song (2015) realizan una revisión sistemática sobre las revisiones sistemáticas en el campo de la Biblioteca y Ciencias del Información, este estudio ha servido como bibliografía en una parte muy importante de la elaboración de este capítulo de la tesis.

I.1.2.1. Objetivos y cuestiones

La búsqueda bibliográfica ha sido uno de los capítulos imprescindibles y más extensos en este trabajo de investigación. La bibliografía documenta el estudio, y pretende ser un sólido fundamento y aval de su pertinencia.

Se ha realizado una revisión sistemática de la literatura sobre los temas de investigación de esta tesis de los que no se han hallado revisiones sistemáticas que se pudiesen replicar. El propósito ha sido obtener una bibliografía basada en la mejor investigación disponible en la que fundamentar el trabajo, tanto para la elaboración del estado del arte como para el desarrollo de la metodología y la práctica. Los temas de investigación sobre los que se necesitaba bibliografía se han sido traducidos a preguntas de investigación. En el presente estudio se han explorado las siguientes cuestiones:

1. ¿El acceso abierto aumenta la visibilidad, la citación y el impacto de autores e instituciones?

2. ¿Existe o no una ventaja de cita en los trabajos de acceso abierto sobre los que no lo son?
3. ¿Qué es la literatura gris científica? ¿Cómo se difunde? ¿Se cita este tipo de literatura?
4. ¿Las tesis doctorales en acceso abierto tienen más visibilidad que la tesis que no son de acceso libre?
5. ¿Las tesis doctorales en acceso abierto son más citadas que las tesis que no son de acceso abierto?
6. ¿La interoperabilidad de los repositorios institucionales aumenta la visibilidad y el impacto?
7. ¿Hay una relación directa entre la interoperabilidad de los repositorios institucionales y el aumento de su visibilidad e impacto?
8. ¿Los aspectos técnicos de los repositorios tienen algo que ver en la visibilidad de estos?
9. ¿Las políticas de mandato de acceso abierto de las instituciones están relacionadas con la visibilidad y el impacto de las mismas?
10. ¿Cómo se miden la visibilidad, citación e impacto del acceso abierto?
11. ¿Qué métricas alternativas se utilizan para medir el impacto?

I.1.3. La revisión de la literatura

En la actualidad se asiste a una gran sobrecarga de información que puede estar desactualizada y cuyas fuentes pueden ser de distinta calidad. Las revisiones de la literatura ofrecen la posibilidad de mantenerse informados en un área de interés, por ello tienen una gran demanda en la mayoría de los campos de investigación. Su necesidad se deriva de la publicación siempre creciente de trabajos científicos (Rapple, 2011). Debido a esa cantidad ingente de trabajos que van apareciendo, los científicos no pueden examinar con detalle cada una de las nuevas publicaciones relevantes para sus intereses (Erren, Cullen, & Erren, 2009). Por ello es ventajoso

y necesario contar con resúmenes regulares de la literatura reciente. Aunque el reconocimiento de los científicos proviene principalmente de la investigación primaria, las revisiones bibliográficas puntuales pueden conducir a nuevos conocimientos sintéticos y son a menudo muy leídas (Hampton & Parker, 2011). Sin embargo, para que estos resúmenes sean útiles tienen que ser compilados de forma profesional (Ketcham & Crawford, 2007).

Maier (2013) hace referencia a que la calidad de la revisión de la literatura es muy importante, ya que su propósito es identificar las lagunas de conocimiento y las necesidades de investigación en el campo y área de investigación del problema en los que el trabajo pretende hacer una contribución al conocimiento. Esto requiere una clara especificación del área del problema y una revisión crítica de la literatura dentro de ese dominio, con el fin de presentar una clara línea argumental que identifique las lagunas de conocimientos y las necesidades de investigación que deben ser abordadas (Maier, 2013). Al igual que las necesidades de investigación y la identificación de las lagunas de conocimiento tienen un efecto directo sobre los objetivos de la investigación, la calidad de la revisión de la literatura también tiene un impacto significativo en la calidad del resto del documento, incluyendo la metodología, resultados y conclusiones, como se ilustra en la Figura 2.

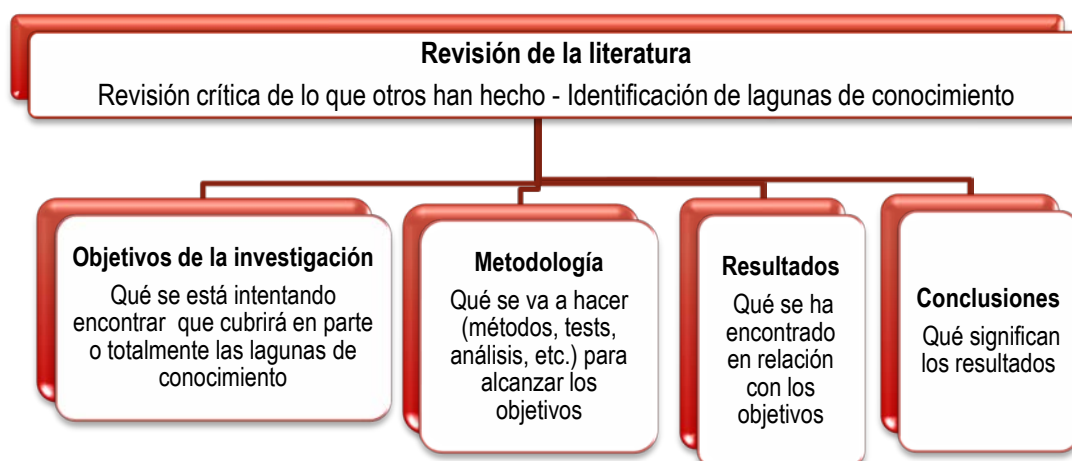


Figura 2. Modificado de Maier (2013). Representación conceptual del impacto de la revisión de la literatura sobre Objetivos, Métodos, Resultados y Conclusiones.

Pautasso (2013) explica que la revisión de la literatura puede llegar a convertirse en un trabajo de titanes si se comienza desde cero. Es por eso que los investigadores que han dedicado su carrera a trabajar en un determinado tema de investigación se encuentran en una posición perfecta para revisar esa bibliografía. Según este autor, se requiere la capacidad de hacer “malabares” con múltiples tareas, como la búsqueda y evaluación de material relacionado con la síntesis de información de diversas fuentes, disponer de pensamiento crítico para parafrasear y evaluar, además de tener destrezas de citación. Este autor define diez reglas simples para escribir una revisión de la literatura (Tabla 2).

Tabla 2

Las 10 reglas simples para escribir una revisión de la literatura

Regla 1	Definir un tópico y una audiencia
Regla 2	Buscar y Re-buscar la literatura
Regla 3	Tomar notas mientras se lee
Regla 4	Escoger el tipo de revisión que se quiere hacer
Regla 5	Mantener la revisión centrada pero que tenga amplio interés
Regla 6	Ser crítico y consistente
Regla 7	Encontrar una estructura lógica
Regla 8	Hacer uso de la retroalimentación
Regla 9	Incluir la propia investigación relevante pero siendo objetivo
Regla 10	Estar al día sin olvidar estudios anteriores

Nota: Basado en Pautasso, M. (2013). Ten Simple Rules for Writing a Literature Review. *PLoS Computational Biology*, 9(7). Fuente: elaboración propia.

I.1.3.1 Tipos de revisiones de la literatura

Según Cooper (1998) existen varios modelos de revisiones que difieren considerablemente en términos de objetivos principales, motivaciones y medios por los que se llevan a cabo (Cooper, 1988). De hecho, el término “revisión” está rodeado de una gran confusión y muchas ambigüedades (Templier & Paré, 2015). Templier y Paré (2015) exponen cómo los

investigadores han utilizado numerosos términos para describir sus métodos y enfoques de revisión, tales como "revisión narrativa" (Joseph, Ng, Koh, & Ang, 2007) "revisión teórica"(Varey, Wood-Harper, & Wood, 2002), "revisión crítica" (Fichman, 1992), "revisión descriptiva"(Yang & Tate, 2009), "revisión exhaustiva"(Liu, Min, & Ji, 2008), "revisión sistemática"(Williams, Dwivedi, Lal, & Schwarz, 2009), y "meta-análisis"(King & He, 2006).

Una revisión sistemática es aquella en la que existe una búsqueda exhaustiva de estudios relevantes sobre un tema. Una vez identificados y obtenidos los estudios, los resultados son sintetizados de acuerdo con un método preestablecido y explícito. Esta forma de revisión ofrece como ventaja la posibilidad de replicarla y verificar si se llega a la misma conclusión (Sáenz, 2001).

Para Petticrew y Roberts, (2006) las revisiones sistemáticas son revisiones de la literatura que están estrechamente unidas a un conjunto de métodos científicos cuyo objetivo es limitar el error sistemático o sesgo, tratando de identificar, evaluar y sintetizar todos los estudios pertinentes con el fin de responder a una pregunta en particular o a un conjunto de preguntas.

Tradicionalmente se han venido realizando un tipo de revisiones llamadas narrativas. Este tipo de revisiones plantean el inconveniente de que utilizan métodos informales, no explícitos y a menudo personales y subjetivos. Los principales inconvenientes de las revisiones narrativas son por un lado que los autores no especifican el proceso que han seguido para buscar, reunir y evaluar la información y, por otro, que debido a la ausencia de esta información, los lectores no son capaces de repetir y verificar los resultados y las conclusiones de las revisiones (Martín Rodero, 2014).

Cynthia Mulrow (1994) describe las revisiones sistemáticas como mecanismos de hipótesis-pruebas. Esto es muy eficaz ya que recuerda al revisor que tiene que enmarcar las preguntas de manera que puedan responderse con claridad y elaborar una pregunta de la revisión como una

hipótesis, o una serie de hipótesis relacionadas, permitiendo que sean enmarcadas claramente, y después aceptadas o rechazadas. Esto no siempre es posible (por ejemplo, para revisiones metodológicas podría resultar más difícil), pero hará que la definición de los criterios de inclusión y exclusión sea más fácil, y ayudará a mantener toda la revisión monitorizada. La regla básica es que nunca se debe comenzar una revisión sistemática hasta que se pueda enmarcar una pregunta con claridad (Mulrow, 1994).

Grant y Booth (2009) describieron 14 tipos de revisiones y las metodologías asociadas a estas. Su propuesta intentaba determinar la terminología prevalente; un proceso conocido como garantía literaria (Research and Innovative Technology Administration (RITA), n.d.). A partir de esto, se identificaron y mapearon los tipos comunes de revisión y sus tópicos clave mediante un marco de Búsqueda, Evaluación, Síntesis y Análisis (*Search, Appraisal, Synthesis and Analysis, SALSA*). Los autores analizaron cada tipo de revisión, describieron sus características y esbozaron las fortalezas y debilidades percibidas (Tabla 3).

Por su parte Petticrew y Roberts (2006) establecen también diversas tipologías de revisiones de la bibliografía. En la Tabla 4 se muestra una síntesis de estos tipos y sus características.

Tabla 3
Tipos de revisiones de la literatura y características

Tipo de revisión	Definición
1. Revisión crítica	Pretende demostrar escritor ha investigado extensamente la literatura y evaluado críticamente su calidad. Va más allá de la mera descripción para incluir grado de análisis e innovación conceptual. Suele generar una hipótesis o modelo
2. Revisión de la literatura	Término genérico: materiales publicados que proporcionan un examen de la literatura reciente o actual. Puede cubrir gran variedad de temas en varios niveles de integridad y exhaustividad. Puede incluir resultados de investigación
3. Revisión de mapeo/ mapeo sistemático	Mapea y categoriza la literatura existente para proyectar otras revisiones y/o investigación mediante la identificación de lagunas en la literatura de investigación
4. Meta-análisis	Técnica que combina estadísticamente los resultados de los estudios cuantitativos para proporcionar un efecto más preciso de los resultados
5. Revisión de estudios mixtos/revisión de métodos mixtos	Se refiere a cualquier combinación de los métodos en los que un componente significativo es una revisión de la literatura (por lo general sistemática). Dentro de un contexto de revisión se refiere a una combinación de enfoques de revisiones, por ejemplo, la combinación de investigación cuantitativa con investigación cualitativa o resultados con estudios de procesos
6. Visión general / Overview	Término genérico: resumen de la literatura que trata de estudiar la literatura y describir sus características
7. Revisión cualitativa sistemática/síntesis de evidencia cualitativa	Método para la integración o la comparación de los resultados de los estudios cualitativos. Busca "temas" o "construcciones" que se encuentran en o entre los estudios cualitativos individuales
8. Revisión rápida	La evaluación de lo que ya se sabe sobre un tema de políticas o prácticas, mediante el uso de métodos de revisión sistemática para buscar y evaluar críticamente la investigación existente
9. <i>Scoping review</i>	Evaluación preliminar del tamaño y del alcance potencial de la literatura de investigación disponible. Pretende identificar la naturaleza y extensión de la evidencia de investigación (normalmente incluida en la investigación en curso)
10. Revisión estado del arte	Tiende a abordar cuestiones más actuales en contraste con otros enfoques combinados retrospectivos y actuales. Puede ofrecer nuevas perspectivas sobre el tema o señalar el área para futuras investigaciones
11. Revisión sistemática	Procura buscar sistemáticamente, evaluar e investigar la síntesis de la evidencia, a menudo adhiriéndose a las directrices sobre cómo llevar a cabo una revisión
12. Revisión sistemática y revisión	Combina las fortalezas de la revisión crítica con un proceso de búsqueda exhaustiva. Aborda normalmente preguntas amplias para producir una "mejor síntesis de la evidencia"
13. Revisión sistematizada	Intenta incluir elementos del proceso de revisión sistemática sin llegar a revisión sistemática. Normalmente realizada como tarea de estudiantes de postgrado
14. Revisión paraguas	Se refiere específicamente a revisar compilando la evidencia desde múltiples revisiones dentro de un documento accesible y utilizable

Nota: basado en Grant y Booth (2009). Fuente: elaboración propia.

Tabla 4

Tipos de revisiones de la literatura y características

Tipo de revisión	Definición
1. Revisión sistemática	Revisión que tiene como objetivo identificar exhaustivamente todos los estudios pertinentes para responder a una pregunta en particular, y evalúa la validez, o solidez, de cada estudio teniendo esto en cuenta al extraer conclusiones
2. Revisión narrativa	A veces se utiliza para referirse a una revisión sistemática que sintetiza los estudios individuales narrativamente (en lugar de a través de un meta-análisis). Esto implica la extracción sistemática, verificación, y resumir narrativamente información sobre sus métodos y resultados
3. Revisión conceptual /síntesis conceptual	Una revisión que pretende sintetizar las áreas del conocimiento conceptual que pueden contribuir a una mejor comprensión de estos temas. Los objetivos de estas síntesis están destinados a proporcionar una visión general de la literatura en un campo determinado, incluyendo las ideas principales, modelos y debates
4. Revisión rápida	Una revisión de la literatura llevada a cabo (a menudo de forma sistemática) pero dentro de un tiempo limitado (a veces semanas o un par de meses) y con restricciones en el ámbito de la búsqueda (por ejemplo, restringidos por año, o país)
5. Revisión realista (o síntesis realista)	Aproximación a la revisión de estudios que implica sintetizar los estudios individuales con el fin de producir teorías generalizables (en lugar de sintetizar los resultados entre los estudios (como las revisiones sistemáticas hacen)
6. Revisión del ámbito	Una revisión del ámbito o alcance de la literatura existente, a veces llevada a cabo antes de una revisión sistemática completa, con el fin de evaluar los tipos de estudios realizados hasta la fecha, y dónde están localizados. Esto puede ayudar con el refinamiento de la pregunta de la revisión completa, y con la estimación de los recursos que se necesitarán
7. Revisión tradicional	Este término a veces se usa para referirse a una revisión de la literatura que no utiliza métodos de revisión sistemática. Estas revisiones todavía pueden representar excelentes visiones generales de la literatura y más amplios los conceptos, no solo comentarios de los resultados
8. Revisión crítica	Este término a veces se utiliza para describir una revisión de la literatura que evalúa una teoría o hipótesis examinando críticamente los métodos y resultados de los estudios primarios, a menudo con una gran cantidad de antecedentes y material de contexto, aunque no con el enfoque formal de una revisión sistemática
9. Revisión experta	Revisión de la literatura, común en la medicina y en ciencias básicas, escrito por un reconocido experto (o grupo de expertos) en el campo
10. Revisión "estado del arte"	Este término se utiliza a veces para referirse a las revisiones diseñadas para tener a los lectores al corriente de las investigaciones más recientes sobre un tema específico. Lo que constituye " reciente " puede variar, al igual que los métodos de las revisiones

Nota: basado Petticrew y Roberts (2006). Fuente: elaboración propia.

I.1.3.2. La revisión sistemática de la literatura en Ciencias Sociales, Humanidades y Bibliotecas y Ciencias de la Información y Documentación

Los bibliotecarios del campo de la medicina se encuentran entre los primeros en adoptar esta metodología y aplicarla al campo de las Bibliotecas y de las Ciencias de la Información (Brettle, 2003). Más tarde, su aplicación se extendió hacia otras áreas no médicas, tales como la académica, la ingeniería del *software* (Kitchenham *et al.*, 2009; Kitchenham & Charters, 2007) o los programas de enseñanza de Bibliotecas (Grant & Booth, 2009).

De la lectura de los trabajos de todos estos autores se deduce que la revisión sistemática deberá cumplir al menos tres condiciones:

1. Utilizar bases de datos académicas como fuente principal.
2. Indicar los criterios de inclusión (y si es el caso, de exclusión) para seleccionar las obras que han formado parte del corpus de análisis.
3. Proporcionar datos para replicar el estudio de revisión.

Respecto a las revisiones sistemáticas en el campo de las Bibliotecas y Ciencias de la Información y Documentación, como se ha mencionado anteriormente hay un importante vacío en la aplicación de este método de revisión. Existen directrices y listas de verificación de la evaluación de la calidad sobre revisiones sistemáticas en general, sin embargo, aparte de los informes sobre la forma en que se utilizan en investigaciones médicas, son pocos los estudios que tratan de establecer criterios adecuados para evaluar la calidad de los estudios y las directrices del proceso, así como la falta de información transparente y completa de las revisiones sistemáticas en este campo. En la mayoría de los casos, esta insuficiencia se debe a la limitada cantidad de detalles metodológicos y a la falta de procesos de revisión rigurosos, explícitos y replicables. Como apuntan Koufogiannakis y Crumley (2006) los métodos de

investigación más rigurosos mejorarían la calidad de la biblioteca de investigación y la literatura de la biblioteca.

El estudio de Xu, Kang y Song (2015) presenta un estado del arte de las revisiones sistemáticas en el campo de las Bibliotecas y las Ciencias de la Información (LIS). Este estudio tiene por objeto proporcionar una visión general del tema. Utilizando los datos en inglés y en chino, los autores intentan investigar exhaustivamente la introducción, discusión y ejecución de los métodos de investigación de la revisión sistemática en LIS. Esto implica examinar la situación actual y la calidad de la metodología empleada en LIS así como proponer un marco de valoración crítica que sea apropiado en este campo de acuerdo a las pautas establecidas.

Además del trabajo de los autores citados, para actualizar y completar la bibliografía sobre el tema de la revisión sistemática de la literatura se ha utilizado el estudio de Templier y Paré, (2015) así como la obra de Petticrew y Roberts (2006).

Esta investigación se ha centrado en realizar una revisión sistemática de la literatura, que de acuerdo a Petticrew y Roberts (2006) ha tenido como objetivo identificar exhaustivamente todos los estudios pertinentes para responder a una pregunta en particular, y evaluar la validez, o solidez, de cada estudio teniéndolos en cuenta al extraer las conclusiones.

I.1.4. El proceso de revisión

Antes de comenzar un proceso de revisión hay que preguntarse y responder a las dos cuestiones que se indican en la Figura 3.

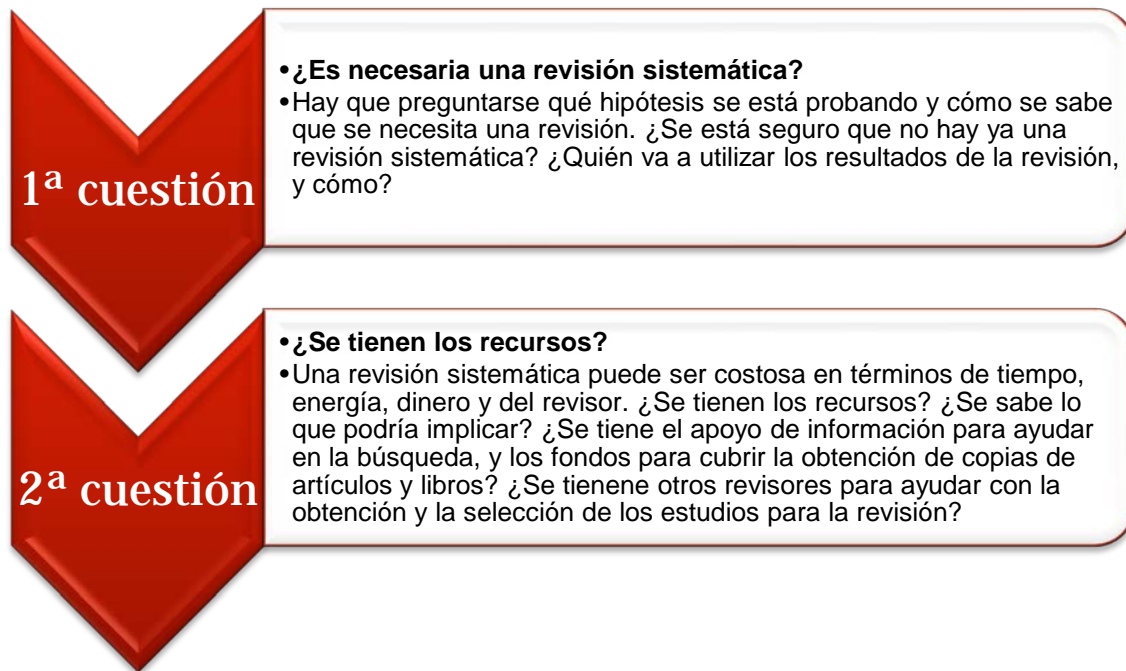


Figura 3. Cuestiones previas a la realización de una revisión sistemática. Fuente: elaboración propia

Si la respuesta es afirmativa para las dos preguntas anteriores será pertinente comenzar el proceso de revisión. En el caso de esta Tesis Doctoral la revisión sistemática es necesaria para fundamentar la literatura del tema de estudio y además se cuenta con los recursos necesarios para poder llevarla a cabo.

I.1.4.1. Método o Protocolo

Como indican Petticrew y Roberts (2006) es esencial escribir un protocolo que indique la pregunta de la revisión, los métodos que se utilizarán, los tipos de estudios y diseños que el revisor tiene la intención de localizar, no solo por qué medios, sino también cómo serán evaluados y sintetizados estos estudios.

El proceso de revisión sistemática se ha llevado a cabo desde agosto de 2014 hasta diciembre de 2015. La metodología empleada se basó en la bibliografía sobre revisiones sistemáticas de la literatura utilizada para la redacción de este capítulo.

I.1.4.1.1. Definir la pregunta. Tipos de preguntas

Al comienzo de la revisión ha sido necesario especificar claramente las preguntas que la revisión tiene como objetivo responder.

Aunque las revisiones sistemáticas a menudo tienen como objetivo responder a preguntas individuales, o probar una sola hipótesis, a veces las cuestiones estratégicas, como es este caso, hacen que el campo de preguntas sea mucho más amplio y puede ser necesario trabajar para identificar cuál es la pregunta o preguntas más probables que será preciso responder.

Concretamente, el primer paso ha sido identificar qué cuestiones era necesario responder y las más obvias pueden estar relacionadas con la relación entre el acceso abierto y el aumento de visibilidad, citación e impacto de autores e instituciones y el papel que desempeñan en este campo los repositorios institucionales. Para determinar si existe una ventaja de cita en los trabajos que están en acceso abierto, a través de los repositorios institucionales, sobre los que no están en abierto se necesita hacer una revisión sistemática de la literatura sobre esta cuestión. Para restringir más el tema al área de la literatura gris científica se requiere realizar una revisión sistemática sobre este tipo de literatura, determinando qué es, qué tipos de documentos se consideran literatura gris, cómo se difunden y qué grado de reutilización y citación tienen en la publicación científica. En el plano de políticas, una revisión sistemática de las intervenciones de las entidades y los organismos de financiación para promover el acceso abierto puede sugerir qué políticas pueden ser efectivas y cuáles adversas sobre el aumento de visibilidad, citación e impacto de la literatura gris científica, y en concreto en las tesis doctorales. En el plano técnico, una revisión sistemática sobre la relación entre interoperabilidad de los repositorios institucionales y entre el aumento de visibilidad de estos puede sugerir también una relación directa de la interoperabilidad de los repositorios con el aumento de citación e impacto de sus contenidos. En el ámbito de la bibliometría es necesaria una revisión sistemática de la literatura

sobre los sistemas de métricas tradicionales y sobre las métricas alternativas eficientes para medir la visibilidad, la citación y el impacto del acceso abierto, en concreto de las tesis doctorales depositadas en los repositorios institucionales.

A partir de estas reflexiones previas a la revisión, el siguiente paso ha sido la redefinición de las cuestiones y de los límites siguiendo las pautas que se describen a continuación. Las cuestiones previas planteadas para realizar las revisiones sistemáticas eran muy generales y con poca profundización en los temas. En un análisis posterior se refinaron estas preguntas reduciéndolas a cuestiones mucho más específicas de tal forma que englobaran a las más genéricas, con el fin de realizar una revisión sistemática mucho más depurada para evitar resultados generales que serían fáciles de obtener con una revisión narrativa de la bibliografía o una revisión del ámbito.

Las cuestiones planteadas finalmente se agruparon en cinco bloques de preguntas relacionados a continuación:

1. ¿El acceso abierto aumenta la visibilidad, la citación y el impacto de autores e instituciones? ¿Existe o no una ventaja de cita en los trabajos de acceso abierto sobre los que no lo son?
2. ¿Qué es la literatura gris científica? ¿Cómo se difunde? ¿Se cita este tipo de literatura? ¿Las tesis doctorales en acceso abierto tienen más visibilidad que la tesis que no son de acceso libre? ¿Las tesis en acceso abierto son más citadas que las tesis que no son de acceso abierto?
3. ¿La interoperabilidad de los repositorios institucionales aumenta la visibilidad y el impacto? ¿Hay una relación directa entre la interoperabilidad de los repositorios institucionales y el aumento de su visibilidad e impacto? ¿Los aspectos técnicos de los repositorios tienen algo que ver en la visibilidad de estos?
4. ¿Las políticas de mandato de las instituciones están relacionadas con la visibilidad y el impacto de las mismas?

5. ¿Cómo se mide la visibilidad, citación e impacto del acceso abierto? ¿Qué métricas alternativas se utilizan para medir el impacto?

Ha sido útil comenzar dividiendo la pregunta de la revisión en sub-preguntas. Si la revisión pretende dar respuesta a una pregunta sobre eficacia (por ejemplo la eficacia del acceso abierto sobre las tesis en el aumento de citación), la pregunta se puede enmarcar utilizando el modelo PICO (Población, Intervención, Comparación, Resultados), que incita al investigador a considerar los componentes de la cuestión de esta forma:

1. Población: ¿En qué población se está interesado? ¿Literatura gris científica? Y si es así ¿qué tipo? ¿Tesis doctorales en acceso abierto?
2. Intervención: Exactamente ¿qué intervención se necesita aplicar? ¿Es una intervención o un conjunto de intervenciones? Puede haber muchas intervenciones relacionadas utilizadas para abordar el problema. Por ejemplo, una intervención sería el depósito de las tesis en los repositorios institucionales, y otras intervenciones relacionadas serían: las políticas de acceso abierto que obligan a depositar las tesis, el grado de interoperabilidad de los repositorios para poder ser recolectados, la medición de la visibilidad y el impacto de las tesis, etc.
3. Comparación: ¿Con qué será comparada la intervención? Por ejemplo en la investigación sobre la efectividad de difundir las tesis a través de repositorios institucionales, ¿con qué se puede comparar esto? ¿Con otra forma de difundir en acceso abierto que no sea a través de repositorios institucionales? ¿Con la no difusión en acceso abierto?
4. Resultados: Existe una amplia gama de resultados para muchas de las intervenciones, la evaluación de la efectividad implica recoger información tanto de impactos positivos como negativos, así como evaluar el equilibrio entre ellos. Especificar la pregunta de la

revisión es importante para determinar qué resultados son los más relevantes para responder a la misma.

5. Contexto: Para las revisiones de algunas intervenciones hay un componente adicional, que debe tenerse en cuenta que es el contexto dentro del cual se desarrolla la intervención.

Tabla 5
Modelo PICOC

Población <i>Population</i> <i>P</i>	Intervención <i>Intervention</i> <i>I</i>	Comparación <i>Comparison</i> <i>C</i>	Resultados <i>Outcomes</i> <i>O</i>	(Contexto) <i>(Context)</i> <i>(C)</i>
Literatura gris científica (Tesis)	Difusión de tesis a través de repositorios institucionales en acceso abierto	Tesis que no están en acceso abierto	Aumento de visibilidad y de impacto de las tesis en acceso abierto	Universidad de Salamanca. 2006-2010
Literatura gris científica (Tesis)	Mandato de acceso abierto las Instituciones a partir de una fecha	Tesis que no estaban sujetas a mandato de esas mismas instituciones	Aumento de visibilidad y de impacto de las tesis en acceso abierto por mandato institucional	Universidad de Salamanca. 2008-2010
Repositorios institucionales	Implementación de herramientas en los repositorios para lograr la interoperabilidad	Comparación entre varios repositorios midiendo la relación entre interoperabilidad y grado de visibilidad	Los repositorios más interoperables aumentan su visibilidad	Ámbito internacional

Nota: Elaboración propia.

I.1.4.1.2. Establecer criterios de inclusión y exclusión

Después de definir las cuestiones de investigación y los objetivos, el siguiente paso en una revisión sistemática de la literatura es definir los criterios de inclusión y exclusión de los estudios.

Los criterios de inclusión y exclusión establecidos se describen a continuación.

Criterios de inclusión:

- Estudios realizados sobre los temas propuestos en las preguntas de investigación que indiquen estado de la cuestión, resultados y conclusiones, y en algunos casos puedan resolver las preguntas más específicas por medio de un proceso de análisis.

- Áreas de conocimiento: los trabajos pueden enmarcarse en cualquier área del conocimiento que trate los temas relacionados con las preguntas de investigación, pero si los resultados de la búsqueda son muy numerosos y no son pertinentes se refinarán por áreas de investigación restringiendo a: *Information Science Library Science, Computer science* y *Communication*.
- Tipos de estudios que se necesitan localizar para responder a las preguntas:
 - Artículos.
 - Libros y capítulos de libros.
 - Tesis.
 - Congresos, conferencias.
 - Revisiones.
 - Bibliografías.
- Lengua de los trabajos: Español, Inglés y Francés.
- Trabajos publicados desde 1998 hasta ahora. Se ha realizado la revisión buscando trabajos publicados a partir de 1998 por ser esta fecha cuando comienza a ser más notorio el movimiento Open Access (“Timeline of the open access movement,” n.d.), tema central de esta investigación.

Criterios de exclusión:

- Contenidos que se aparten de los temas de las cuestiones planteadas.
- Otros tipos de estudios.
- Otras lenguas.
- Trabajos publicados antes de 1998.
- Trabajos a los que no estén en acceso abierto o a los que no tengamos acceso a través de suscripción de la USAL.

I.1.4.1.3. Llevar a cabo la búsqueda de literatura. Estrategia de búsqueda para la identificación de la literatura

Después de haber decidido las cuestiones de investigación y los criterios de inclusión y exclusión, el siguiente paso es llevar a cabo una búsqueda exhaustiva de la literatura para localizar los estudios pertinentes. Identificar y localizar las evidencias o pruebas disponibles en la literatura implica establecer una estrategia de búsqueda bibliográfica sistemática que permita hallar la literatura relevante, publicada o no, que contenga posibles respuestas a las interrogantes establecidas. Se precisa, además, que esta sea fácilmente reproducible y aplicable a la mayoría de las bases de datos existentes y a otros recursos disponibles. La estrategia deberá basarse en los elementos de la pregunta de revisión junto con el diseño del estudio que se considere más adecuado (Wanden-Berghe y Sanz-Valero, 2014). Esto implica determinar la localización (dónde) y la terminología (cómo) que se va a emplear en la búsqueda. Es imprescindible conocer las fuentes de información disponibles para dar respuesta a las preguntas formuladas y lograr una eficiente recuperación de la información.

Según Petticrew y Roberts (2006) hay muchas formas de información a tener en cuenta, pero la mayoría de las veces el punto de partida es la búsqueda en bases de datos electrónicas. Sin embargo, no solo las bases de datos son la única fuente de la literatura, y a veces no son las más útiles. En muchas áreas de investigación, particularmente en las ciencias sociales el grueso de la evidencia relevante podría no aparecer en revistas sino localizarse en informes de literatura gris que no estarían indexados en bases de datos electrónicas. El tipo de información solicitada dependerá tanto de la pregunta de revisión como de la inclusión de los criterios. Se incluirán artículos publicados en revistas, y otros informes y estudios publicados o no. Estos últimos podrían localizarse a través de búsquedas de ponencias en conferencias, resúmenes de tesis y otra literatura gris. Las búsquedas también necesitan incluir capítulos de libros, sobre todo en

ciencias sociales, porque estos podrían hacer referencia a resultados de estudios nunca publicados en revisiones (sistemáticas o no) y también podrían buscarse a mano estudios primarios. A todos estos criterios expuestos por Petticrew y Roberts (2006) con los que se está de acuerdo, se ha añadido la búsqueda en recursos propios del acceso abierto: repositorios, recolectores, bibliotecas digitales y buscadores científicos como *Google Scholar*.

En cuanto a la terminología y formulación de la estrategia de búsqueda, una ecuación de búsqueda eficaz sería, por ejemplo, la formada por descriptores y sus correspondientes calificadores combinados entre sí mediante los operadores booleanos más apropiados. La utilización de descriptores es una opción adecuada para localizar los trabajos relacionados con un tema de interés, para facilitar su recuperación y para dar visibilidad a los artículos científicos. Los descriptores no solo son útiles para realizar una búsqueda bibliográfica, sino que además sirven para analizar los trabajos por áreas de conocimiento (Wanden-Berghe y Sanz-Valero, 2014). Para realizar las búsquedas se han utilizado tanto descriptores como palabras clave.

Antes de formular una ecuación de búsqueda se deben establecer de manera clara y precisa los términos de búsqueda para a continuación establecer las relaciones lógicas que se darán entre ellos. Para unir conceptos relacionados se utilizará el operador OR (operador de unión); para relacionar términos que hacen referencia a conceptos distintos en un mismo documento se utilizará AND (operador de intersección); para eliminar documentos que contenga el término no deseado se empleará NOT (operador de exclusión). A la hora de formular ecuaciones de búsqueda más complejas, en las que se combinen varios operadores, se utilizarán paréntesis para indicar qué operación se debe efectuar en primer lugar. (Martín Rodero, 2014).

Al comenzar esta investigación se cuestionado por dónde empezar. De acuerdo a Petticrew y Roberts (2006) no tiene sentido hacer una revisión sistemática que ya se haya hecho antes, a no ser que sea evidente que las anteriores revisiones sistemáticas están sesgadas, o están

significativamente anticuadas, por ejemplo, si se han publicado nuevos estudios desde que se completó la revisión existente. Por esta razón, es importante comenzar cualquier nueva revisión mediante la búsqueda de revisiones sistemáticas existentes, para evitar la duplicación.

Por tanto, el primer paso que se ha dado en la estrategia de búsqueda ha sido localizar revisiones sistemáticas de la literatura generales en el campo de las Bibliotecas y las Ciencias de la Información por un lado y, por otro, localizar revisiones sistemáticas de la literatura específicas sobre las preguntas de investigación.

La búsqueda de las revisiones sistemáticas se realizó mediante métodos de recuperación bibliográfica habituales, seleccionándose aquellos artículos accesibles a texto completo. Se incluyeron los estudios redactados en español, inglés, y francés.

Se ha utilizado la revisión sistemática de Xu, Kang y Song (2015) para localizar la bibliografía sobre “revisiones sistemáticas de la literatura en LIS”, y se ha utilizado el estudio de estos autores para la redacción de este capítulo.

De los temas generales como el acceso abierto, los repositorios, la literatura gris científica y la bibliometría no se ha realizado revisión sistemática de la literatura sino que se ha realizado lo que Petricrew y Roberts (2006) denominan una revisión del ámbito o alcance de la literatura existente, a veces llevada a cabo antes de una revisión sistemática completa, con el fin de evaluar los tipos de estudios realizados sobre estos temas hasta la fecha, y dónde están localizados. Esto ha ayudado con el refinamiento de las preguntas de la revisión sistemática completa, y con la estimación de los recursos que se han necesitado.

Para realizar la revisión del ámbito se han buscado y localizado bibliografías sobre estos temas o se han realizado búsquedas a mano por palabra clave o autores de relevancia. Esta búsqueda se ha realizado fundamentalmente en *Google Scholar* que indiza todos los recursos en acceso abierto que interesaba rastrear: repositorios, recolectores y revistas de acceso abierto.

Se han empleado estrategias de búsqueda exhaustivas de acuerdo a las ecuaciones de búsqueda establecidas a partir de las preguntas de investigación con el fin de identificar toda la literatura relevante, se ha comenzado buscando en *Web of Science (WOS)*, en la colección principal, desde 1998 hasta 2015.

Las otras bases de datos elegidas han sido *Library and information science abstracts (LISA)*, *Library, Information Science and Technology Abstracts (LISTA)* y *Scopus*. Las estrategias de búsqueda para LISA (Proquest), LISTA (Ebscohost) y Scopus (Elsevier) se basaron en la estructura de indexación de cada base de datos para llevar a cabo consultas similares a la utilizada en WOS con un rango de fechas de 1998 hasta 2015. Las búsquedas se restringieron a las publicaciones en inglés, español y francés. Se obtuvieron un total de 2.192 referencias por medio de diversas estrategias de búsqueda, y los resultados de la búsqueda fueron descargados a *Mendeley* donde fueron eliminados los duplicados y almacenadas las referencias seleccionadas junto con los ficheros pdf del texto completo correspondiente. Además de *Mendeley*, utilizado como gestor de referencias, se utilizó *Excel* como sistema de gestión de la base de datos obtenida.

Se identificaron además citas adicionales a través de métodos suplementarios: se incluyeron citas relevantes de listas de referencias o bibliografías; se realizó el rastreo de citas de todos los trabajos incluidos; se buscaron trabajos de autores expertos en las materias; se hicieron búsquedas puntuales en DOAJ, BASE, TDR, DART-Europe, SCIELO. Las ecuaciones de búsqueda llevadas a cabo en las cuatro bases de datos se replicaron en Google Scholar y se obtuvieron 883 resultados, una vez gestionados en Mendeley y eliminados los duplicados se incluyeron 182 trabajos en la selección de la literatura. El proceso y resultados totales puede verse en el Diagrama de flujo de la selección de estudios para la inclusión en la revisión sistemática que aparece en la Figura 4.

Tabla 6

Otros recursos utilizados para localizar la literatura

Bases de datos	Repositorios/ recolectores	Portales científicos	Buscadores	Bibliografía Revisiones	Blogs y Wikis
WOS	DOAJ	G Scholar	Google	C.W. Bailey	http://scholarlyoa.com http://lis-systematic-reviews.wikispaces.com/Welcome , http://oad.simmons.edu/oadwiki/
SCOPUS	BASE	Mendeley			Blog de COAR
LISA	TDR				OpenScience Blog
LISTA	DART-Europe				OAW
CSIC	NLDTD				OpenAire Blog
	SCIELO				SPARC Blog http://poeticeconomics.blogspot.com.es

Nota: Elaboración propia.

Ecuaciones de búsqueda

A continuación se detallan las ecuaciones de búsqueda que se han realizado de acuerdo a las cuestiones planteadas previamente a la búsqueda bibliográfica. Las ecuaciones se han replicado en las diferentes bases de datos y buscadores a través de una búsqueda avanzada. Las palabras clave utilizadas se han buscado en los campos de materia, en el título del trabajo, y a veces en el resumen. Se ha aplicado un filtro temporal limitando las fechas de publicación de los trabajos al periodo: 1998 – hasta el presente. Además, otro filtro utilizado para la selección final de la bibliografía ha sido la priorización de los resultados de los trabajos que están en acceso abierto.

En las tablas se expresan las búsquedas realizadas en cada una de las bases de datos y el número de resultados obtenidos. Se puede observar que hay ligeras variaciones en las ecuaciones de búsqueda al replicar las mismas preguntas de investigación en cada una de las bases de datos con el fin de adaptarlas a las características de estas.

Para responder al primer bloque de cuestiones -¿El acceso abierto aumenta la visibilidad, la citación y el impacto de autores e instituciones? ¿Existe o no una ventaja de cita en los trabajos de acceso abierto sobre los que no lo son?- se ha formulado la siguiente ecuación de búsqueda: ("Open Access" OR "Acceso abierto") AND (visibili* OR impact* OR cita*). Las búsquedas y los resultados obtenidos en cada base de datos se expresan en las Tablas 7, 8, 9 y 10.

Tabla 7

Resultados obtenidos de WOS a la primera cuestión

WEB OF SCIENCE - FECYT

	Resultados	Búsqueda
#1	49	Título: (("Open Access" OR "Acceso abierto") AND (visibili*OR impact* OR cita*)) AND Idioma: (English OR French OR Spanish) Índices=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI, IC Período de tiempo=1998-2015
#2	194	Tema: (("Open Access" OR "Acceso abierto") AND (visibili*OR impact* OR cita*)) AND Idioma: (English OR French OR Spanish) Refinado por: Categorías de Web of Science: (INFORMATION SCIENCE LIBRARY SCIENCE OR COMPUTER SCIENCE INFORMATION SYSTEMS OR COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS OR MULTIDISCIPLINARY SCIENCES) Período de tiempo=1998-2015
#3	211	#1 OR #2

Nota: Elaboración propia

Tabla 8

Resultados obtenidos en Scopus a la primera cuestión

SCOPUS

	Resultados	Búsqueda
#1	159	TITLE (("Open Access" OR "Acceso abierto") AND (visibili* OR impact* OR cita*)) AND PUBYEAR > 1997 AND PUBYEAR < 2016 AND (LIMIT-TO (LANGUAGE, "English") OR LIMIT-TO (LANGUAGE, "Spanish"))
#2	124	KEY (("Open Access" OR "Acceso abierto") AND (visibili* OR impact* OR cita*)) AND PUBYEAR > 1997 AND PUBYEAR < 2016 AND (LIMIT-TO (LANGUAGE, "English") OR LIMIT-TO (LANGUAGE, "Spanish") OR LIMIT-TO (LANGUAGE, "French")) AND (LIMIT-TO (SUBJAREA, "SOC") OR LIMIT-TO (SUBJAREA, "COMP") OR LIMIT-TO (SUBJAREA, "MULT"))
#3	308	#1 OR #2

Nota: Elaboración propia

Tabla 9

Resultados obtenidos en LISTA a la primera cuestión

LISTA

	Resultados	Búsqueda
# 1	64	TI ("Open Access" OR "Acceso abierto") AND (visibili* OR impact* OR cita*) Limiters: - Publication Date: 19980101-20151231 - Language: English, Spanish, French
# 2	74	SU("Open Access" OR "Acceso abierto") AND (visibili* OR impact* OR cita*) Limiters: - Publication Date: 19980101-20151231. - Language: English, Spanish, French
#3	52	KW ("Open Access" OR "Acceso abierto") AND (visibili* OR impact* OR cita*) Limiters: - Publication Date: 19980101-20151231. - Language: English, Spanish, French
#3	60	#1 OR #2 OR #3

Nota: Elaboración propia

Tabla 10

Resultados obtenidos en LISA a la primera cuestión

LISA

	Resultados	Búsqueda
# 1	62	ti(("Open Access" OR "Acceso abierto") AND (visibili* OR impact* OR cita*)) Límites adicionales - Fecha: Desde 1998 hasta 2015; Idioma: Español, Francés, Inglés
# 2	96	su(("Open Access" OR "Acceso abierto") AND (visibili* OR impact* OR cita*)) Límites adicionales - Fecha: Desde 1998 hasta 2015; Idioma: Español, Francés, Inglés
#3	96	#1 OR #2 OR #3

Nota: Elaboración propia

El segundo bloque de cuestiones está relacionado con la literatura gris científica y las tesis doctorales -¿Qué es la literatura gris científica? ¿Cómo se difunde? ¿Se cita este tipo de literatura? ¿Las tesis doctorales en acceso abierto tienen más visibilidad que la tesis que no son de acceso libre? ¿Las tesis en acceso abierto son más citadas que las tesis que no son de acceso abierto?- y las ecuaciones de búsqueda para realizar la revisión sistemática sobre estas cuestiones han sido las siguientes: ("literatura gris" OR "grey literature" OR "gray literature" OR "littérature grise" OR e-theses OR theses OR dissertations OR tesis OR "tesis electrónicas") AND ((dissemination OR diffusion OR difusión OR diseminación) OR (citation OR citación)).

Las Tablas 11, 12, 13 y 14 muestran los resultados obtenidos en cada una de las bases de datos.

Tabla 11

Búsqueda y resultados en WOS de la segunda cuestión

WEB OF SCIENCE (WOS)

	Resultados	Búsqueda
# 1	52	Título: (((("literatura gris" OR "grey literature" OR "gray literature" OR "littérature grise" OR e-theses OR theses OR dissertations OR tesis OR "tesis electrónicas") AND ((dissemination OR diffusion OR difusión OR diseminación) OR (citation OR citación)))) Período de tiempo=1998-2015
#2	197	Tema: (((("literatura gris" OR "grey literature" OR "gray literature" OR "littérature grise" OR e-theses OR theses OR dissertations OR tesis OR "tesis electrónicas") AND ((dissemination OR diffusion OR difusión OR diseminación) OR (citation OR citación)))) Refinado por: Categorías de Web of Science: (INFORMATION SCIENCE LIBRARY SCIENCE OR COMPUTER SCIENCE INFORMATION SYSTEMS OR SOCIAL SCIENCES INTERDISCIPLINARY) Índices=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI, CCR-EXPANDED, IC Período de tiempo=1998-2015
#3	227	#1 OR #2

Nota: Elaboración propia

Tabla 12

Búsqueda y resultados en SCOPUS de la segunda cuestión

SCOPUS

	Resultados	Búsqueda
# 1	81	TITLE (("literatura gris" OR "grey literature" OR "gray literature" OR "littérature grise" OR e-theses OR theses OR dissertations OR tesis OR "tesis electrónicas")) AND TITLE (((dissemination OR diffusion OR difusión OR diseminación) OR (citation OR citación))))
#2	105	TITLE-ABS-KEY (((("literatura gris" OR "grey literature" OR "gray literature" OR "littérature grise" OR e-theses OR theses OR dissertations OR tesis OR "tesis electrónicas") AND ((dissemination OR diffusion OR difusión OR diseminación) OR (citation OR citación))))) AND PUBYEAR > 1997
#3	186	#1 OR #2

Nota: Elaboración propia

Tabla 13

Búsqueda y resultados en LISTA de la segunda cuestión

LISTA

	Resultados	Búsqueda
# 1	78	TI= (((("literatura gris" OR "grey literature" OR "gray literature" OR "littérature grise" OR e-theses OR theses OR dissertations OR tesis OR "tesis electrónicas") AND ((dissemination OR diffusion OR difusión OR diseminación) OR (citation OR citación))))
#2	110	SU= (((("literatura gris" OR "grey literature" OR "gray literature" OR "littérature grise" OR e-theses OR theses OR dissertations OR tesis OR "tesis electrónicas") AND ((dissemination OR diffusion OR difusión OR diseminación) OR (citation OR citación))))
#3	110	#1 OR #2

Nota: Elaboración propia

Tabla 14

Búsqueda y resultados en LISA de la segunda cuestión

LISA

	Resultados	Búsqueda
# 1	86	ti(("literatura gris" OR "grey literature" OR "gray literature" OR "littérature grise" OR e-theses OR theses OR dissertations OR tesis OR "tesis electrónicas") AND ((dissemination OR diffusion OR difusión OR diseminación) OR (citation OR citación)))
#2	141	SU= (((("literatura gris" OR "grey literature" OR "gray literature" OR "littérature grise" OR e-theses OR theses OR dissertations OR tesis OR "tesis electrónicas") AND ((dissemination OR diffusion OR difusión OR diseminación) OR (citation OR citación))))
#3	137	#1 OR #2

Nota: Elaboración propia

El bloque 3 trata sobre los aspectos técnicos relacionados con la interoperabilidad de los repositorios y la posible influencia de estos aspectos en el aumento de la visibilidad y el impacto: ¿La interoperabilidad de los repositorios institucionales aumenta la visibilidad y el impacto? ¿Hay una relación directa entre la interoperabilidad de los repositorios institucionales y el aumento de su visibilidad e impacto? ¿Los aspectos técnicos de los repositorios tienen algo que ver en la visibilidad de estos? La ecuación de búsqueda formulada es la siguiente: Repositor* AND Interoperabili*) AND (visibili* OR impact*). Las Tablas 15, 16, 17 y 18 muestran los resultados de la búsqueda en las cuatro bases de datos.

Tabla 15

Búsqueda y resultados en WOS de la tercera cuestión

WEB OF SCIENCE (WOS)

	Resultados	Búsqueda
# 1	0	Título: (Repositor* AND Interoperabili*) AND (visibility* OR impact *) Período de tiempo=1998-2015
# 2	21	Título: Repositor* AND Interoperabili*. Período de tiempo=1998-2015
#3	31	Tema: (Repositor* AND (Interoperabili*) AND (visibility* OR impact *) Período de tiempo=1998-2015
#4	51	#1 OR #2 OR #3

Nota: Elaboración propia

Tabla 16

Búsqueda y resultados en SCOPUS de la tercera cuestión

SCOPUS

	Resultados	Búsqueda
# 1	0	TITLE ((repositor* AND interoperabili*) AND (visibility* OR impact *)) AND PUBYEAR > 1997
#2	46	TITLE (Repositor* AND Interoperabili*) AND PUBYEAR > 1997
#3	55	TITLE-ABS-KEY ((repositor* AND (interoperability OR interoperabilidad OR interoperabilité) AND (visibility OR visibilidad OR visibilité OR impact OR impacto))) AND PUBYEAR > 1997
#4	97	#1 OR #2 OR #3

Nota: Elaboración propia

Tabla 17

Búsqueda y resultados en LISTA de la tercera cuestión

LISTA

	Resultados	Búsqueda
# 1	0	TI= (repositor* AND interoperabili*) AND (visibili* OR impact*). Límites: 1998-2015. English, French, Spanish
#2	12	TI= repositor* AND interoperabili*. Límites: 1998-2015. English, French, Spanish
#3	0	SU-KW= TI= (repositor* AND interoperabili*) AND (visibili* OR impact*). Límites: 1998-2015. English, French, Spanish
#4	5	AB= (repositor* AND interoperabili*) AND (visibili* OR impact*). Límites: 1998-2015. English, French, Spanish
#5	14	#1 OR #2 OR #3 OR #4

Nota: Elaboración propia

Tabla 18

Búsqueda y resultados en LISA de la tercera cuestión

LISA

	Resultados	Búsqueda
# 1	0	ti((repositor* AND interoperabili*) AND (visibili* OR impact*)) Limitado por: Fecha: Desde 1998 hasta 2015 Idioma: Español, Francés, Inglés
# 2	1	su((repositor* AND interoperabili*) AND (visibili* OR impact*)) Limitado por: Fecha: Desde 1998 hasta 2015 Idioma: Español, Francés, Inglés
#3	7	ab((repositor* AND interoperabili*) AND (visibili* OR impact*)) Limitado por: Fecha: Desde 1998 hasta 2015 Idioma: Español, Francés, Inglés
#4	7	#1 OR #2 OR #3

Nota: Elaboración propia

Las políticas de mandato de acceso abierto son otro de los aspectos que se trata de averiguar si influyen en el aumento de visibilidad e impacto del acceso abierto. Para revisar si existen estudios sobre la cuestión planteada en el bloque 4 -¿Las políticas de mandato de las instituciones están relacionadas con la visibilidad y el impacto de las mismas?- se ha formulado la siguiente ecuación: ("Open Access" OR "Acceso abierto") AND (mandat* OR poli*) AND (visibili* OR impact*). Los resultados se muestran en las Tablas 19, 20, 21 y 22.

Tabla 19

Búsqueda y resultados en WOS de la cuarta cuestión

WEB OF SCIENCE (WOS)

	Resultados	Búsqueda
# 1	17	Título:(("open access" OR "Acceso abierto") AND (mandat* OR poli*)) Refinado por: Categorías de Web of Science: (INFORMATION SCIENCE LIBRARY SCIENCE OR MULTIDISCIPLINARY SCIENCES) Período de tiempo: 1998-2015.
#2	12	Título: (("open access" OR "Acceso abierto") AND (mandat* OR poli*) AND (visibili* OR impact*)) Período de tiempo=1998-2015
	69	Tema: (("open access" OR "Acceso abierto") AND (mandat* OR poli*) AND (visibili* OR impact*)). Refinado por: Categorías de Web of Science: (INFORMATION SCIENCE LIBRARY SCIENCE OR COMPUTER SCIENCE INFORMATION SYSTEMS OR SOCIAL SCIENCES INTERDISCIPLINARY) Período de tiempo: 1998-2015.
#3	89	#1 OR #2 OR #3

Nota: Elaboración propia

Tabla 20

Búsqueda y resultados en SCOPUS de la cuarta cuestión

SCOPUS

	Resultados	Búsqueda
# 1	3	(TITLE (("open access" OR "Acceso abierto")) AND TITLE (mandat*) OR TITLE (poli*) AND TITLE ((impact* OR visibili*))) AND PUBYEAR > 1997 AND PUBYEAR < 2016
#2	33	(TITLE-ABS-KEY (("open access" OR "Acceso abierto")) AND TITLE-ABS-KEY ((mandate* OR mandato*)) AND TITLE-ABS-KEY ((impact* OR visibili*))) AND SUBJAREA (mult OR ceng OR CHEM OR comp OR eart OR ener OR engi OR envi OR mate OR math OR phys OR mult OR arts OR busi OR deci OR econ OR psyc OR soci) AND PUBYEAR > 1997 AND PUBYEAR < 2016
#3	34	#1 OR #2

Nota: Elaboración propia

Tabla 21

Búsqueda y resultados en LISTA de la cuarta cuestión

LISTA

	Resultados	Búsqueda
# 1	0	TI= ("open access" OR "Acceso abierto") AND (mandat* OR poli*) AND (impact* OR visibili*). 1998-2015. Inglés, español, francés.
#2	97	SU-KW-AB ("open access" OR "Acceso abierto") AND (mandat* OR poli*) AND (impact* OR visibili*). 1998-2015. Inglés, español, francés.
#3	97	#1 OR #2

Nota: Elaboración propia

Tabla 22

Búsqueda y resultados en LISA de la cuarta cuestión

LISA

	Resultados	Búsqueda
# 1	0	ti ("open access" OR "Acceso abierto") AND (mandat* OR poli*) AND (impact* OR visibili*) Limitado por: Fecha: Desde 1998 hasta 2015 Idioma: Español, Francés, Inglés
# 2	3	su ("open access" OR "Acceso abierto") AND (mandat* OR poli*) AND (impact* OR visibili*) Limitado por: Fecha: Desde 1998 hasta 2015 Idioma: Español, Francés, Inglés
#3	86	ab ("open access" OR "Acceso abierto") AND (mandat* OR poli*) AND (impact* OR visibili*) Limitado por: Fecha: Desde 1998 hasta 2015 Idioma: Español, Francés, Inglés
#4	83	#1 OR #2 OR #3

Nota: Elaboración propia

Finalmente se trató de obtener información sobre las métricas del acceso abierto, tanto las métricas clásicas como las métricas alternativas que ayudan a medir la visibilidad, la citación y el

impacto del acceso abierto: ¿Cómo se mide la visibilidad, citación e impacto del acceso abierto?
 ¿Qué métricas alternativas se utilizan para medir el impacto?

Ecuaciones de búsqueda: ("Open Access" OR "acceso abierto") AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics); ("open access" AND impact) AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics). Tablas 23-26.

Tabla 23

Búsqueda y resultados en WOS de la quinta cuestión

WEB OF SCIENCE (WOS)

	Resultados	Búsqueda
# 1	8	Título: ("Open Access" OR "acceso abierto") AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics) =1998-2015. Inglés, español, francés.
# 2	1	Título: Impact* AND ("Open Access" OR "acceso abierto") AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics) =1998-2015. Inglés, español, francés.
#3	111	Tema: ("Open Access" OR "acceso abierto") AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics) =1998-2015. Inglés, español, francés.
#4	70	Tema: Impact* AND ("Open Access" OR "acceso abierto") AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics) =1998-2015. Inglés, español, francés.
#5	111	#1 OR #2 OR #3 OR #4

Nota: Elaboración propia

Tabla 24

Búsqueda y resultados en SCOPUS de la quinta cuestión

SCOPUS

	Resultados	Búsqueda
# 1	12	TITLE ("Open Access" OR "acceso abierto") AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics) AND PUBYEAR > 1997. Inglés, español, francés.
#2	1	TITLE : Impact* AND ("Open Access" OR "acceso abierto") AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics) AND PUBYEAR > 1997. Inglés, español, francés.
#3	184	TITLE-ABS-KEY ("Open Access" OR "acceso abierto") AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics) AND PUBYEAR > 1997 . Inglés, español, francés.
	95	TITLE-ABS-KEY Impact* AND ("Open Access" OR "acceso abierto") AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics) AND PUBYEAR > 1997 . Inglés, español, francés.
#4	184	#1 OR #2 OR #3 OR #4

Nota: Elaboración propia

Tabla 25

Búsqueda y resultados en LISTA de la quinta cuestión

LISTA

	Resultados	Búsqueda
# 1	7	TI= ("Open Access" OR "acceso abierto) AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics) . Límites: 1998-2015. English, French, Spanish
#2	0	TI= Impact* AND ("Open Access" OR "acceso abierto) AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics). Límites: 1998-2015. English, French, Spanish
#3	23-17	SU-KW= ("Open Access" OR "acceso abierto) AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics). Límites: 1998-2015. English, French, Spanish
#4	2-0	SU-KW= Impact* AND ("Open Access" OR "acceso abierto) AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics). Límites: 1998-2015. English, French, Spanish
#5	14	#1 OR #2 OR #3 OR #4

Nota: Elaboración propia

Tabla 26

Búsqueda y resultados en LISA de la quinta cuestión

LISA

	Resultados	Búsqueda
# 1	6	ti((Open Access OR acceso abierto) AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics))Limitado por: Fecha: Desde 1998 hasta 2015 Idioma: Español, Francés, Inglés
# 2	73	su((Open Access OR acceso abierto) AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics))Limitado por: Fecha: Desde 1998 hasta 2015 Idioma: Español, Francés, Inglés
	14	su(Impact* AND (Open Access OR acceso abierto) AND (bibliometric* OR almetric* OR informetric* OR scientometric* OR webometrics))Limitado por: Fecha: Desde 1998 hasta 2015 Idioma: Español, Francés, Inglés
#4	76	#1 OR #2 OR #3

Nota: Elaboración propia

Ecuaciones de búsqueda en Google Scholar

La forma de buscar mediante la búsqueda avanzada en Google Scholar difiere de la forma de buscar en las bases de datos anteriores. Por ello, no se han podido replicar exactamente en Google Scholar las ecuaciones de búsqueda utilizadas en las bases de datos y ha sido necesario adaptarlas a las características de este buscador.

Se han realizado las búsquedas por palabras clave exclusivamente en el campo "título del artículo" con el fin de restringir el número de resultados. También se ha especificado el intervalo temporal a 1998-2015.

Bloques de cuestiones, ecuaciones de búsqueda y resultados en Google Scholar

1. ¿El acceso abierto aumenta la visibilidad, la citación y el impacto de autores e instituciones?
¿Existe o no una ventaja de cita en los trabajos de acceso abierto sobre los que no lo son?

allintitle: "open access" visibility OR impact OR citation – 346 resultados

https://scholar.google.com/scholar?as_vis=1&q=allintitle:+%22open+access%22+visibility+OR+impact+OR+citation&hl=es&as_sdt=1,5&as_ylo=1998&as_yhi=2015

allintitle: "acceso abierto" Visibilidad OR impacto OR citación – 28 resultados

https://scholar.google.com/scholar?as_vis=1&q=allintitle:+%22acceso+abierto%22+Visibilidad+OR+impacto+OR+citaci%C3%B3n&hl=es&as_sdt=1,5&as_ylo=1998&as_yhi=2015

No hay resultados representativos para la ecuación de búsqueda en idioma francés.

2. ¿Qué es la literatura gris científica? ¿Cómo se difunde? ¿Se cita este tipo de literatura? ¿Las tesis doctorales en acceso abierto tienen más visibilidad que la tesis que no son de acceso libre?
¿Las tesis en acceso abierto son más citadas que las tesis que no son de acceso abierto?

allintitle: difusión OR diseminación OR citación "literatura gris" – 2 resultados

https://scholar.google.com/scholar?as_q=&as_epq=literatura+gris&as_oq=difusi%C3%B3n+diseminaci%C3%B3n+citaci%C3%B3n&as_eq=&as_occt=title&as_sauthors=&as_publication=&as_ylo=&as_yhi=&btnG=&hl=es&as_sdt=0%2C5

allintitle: "gray literature" OR "grey literature" dissemination OR diffusion OR citation - 39 resultados

https://scholar.google.com/scholar?q=allintitle:+%E2%80%9Cgray+literature%E2%80%9D+OR+%22grey+literature%22+dissemination+OR+diffusion+OR+citation&hl=es&as_sdt=0,5

No hay resultados para la ecuación de búsqueda en idioma francés.

3. ¿La interoperabilidad de los repositorios institucionales aumenta la visibilidad y el impacto?

¿Hay una relación directa entre la interoperabilidad de los repositorios institucionales y el aumento de su visibilidad e impacto? ¿Los aspectos técnicos de los repositorios tienen algo que ver en la visibilidad de estos?

allintitle: Repositories Interoperability OR visibility OR impact – 149 resultados

https://scholar.google.com/scholar?as_vis=1&q=allintitle:+Repositories+Interoperability+OR+visibility+OR+impact&hl=es&as_sdt=1,5&as_ylo=1998&as_yhi=2015

allintitle: Repository Interoperability OR visibility OR impact – 210 resultados

https://scholar.google.com/scholar?as_sdt=1,5&q=allintitle:+Repository+Interoperability+OR+visibility+OR+impact&hl=es&as_ylo=1998&as_yhi=2015&as_vis=1

allintitle: Repositorios Interoperabilidad OR visibilidad OR impacto – 33 resultados

https://scholar.google.com/scholar?q=allintitle%3A+Repositorios+Interoperabilidad+OR+visibilidad+OR+impacto&btnG=&hl=es&as_sdt=1%2C5&as_ylo=1998&as_yhi=2015&as_vis=1

4. ¿Las políticas de mandato de las instituciones están relacionadas con la visibilidad y el impacto de las mismas?

allintitle: "open access" mandate OR politics – 47 resultados

https://scholar.google.com/scholar?as_vis=1&q=allintitle:+%22open+access%22+mandate+OR+politics&hl=es&as_sdt=1,5&as_ylo=1998&as_yhi=2015

allintitle: acceso abierto políticas OR mandato – 25 resultados

https://scholar.google.com/scholar?as_q=acceso+abierto&as_epq=&as_oq=pol%C3%ADticas+mandato&as_eq=&as_occt=title&as_sauthors=&as_publication=&as_ylo=1998&as_yhi=2015&btnG=&hl=es&as_sdt=1%2C5&as_vis=1

5. ¿Cómo se mide la visibilidad, citación e impacto del acceso abierto? ¿Qué métricas alternativas se utilizan para medir el impacto?

allintitle: "open access" bibliometrics OR almetrics OR informetrics OR scientometrics OR webometrics – 4 resultados

https://scholar.google.es/scholar?as_vis=1&q=allintitle:%22open+access%22+bibliometrics+OR+almetrics+OR+informetrics+OR+scientometrics+OR+webometrics&hl=es&as_sdt=1,5&as_ylo=1998&as_yhi=2015

Tabla 27

Tabla Excel de las Búsquedas en Google Scholar

1. ¿El acceso abierto aumenta la visibilidad, la citación y el impacto de autores e instituciones? ¿Existe o no una ventaja de cita en los trabajos de acceso abierto sobre los que no lo son?	Ecuaciones	Nº Resultados	Resultados	
	allintitle: "open access" visibility OR impact OR citation	346	https://scholar.google.com/scholar?as_vis=1&q=allintitle:%22open+access%22+visib	108
	allintitle: "acceso abierto" Visibilidad OR impacto OR citación	28	https://scholar.google.com/scholar?as_vis=1&q=allintitle:%22acceso+abierto%22+vi	19
2. ¿Qué es la literatura gris científica? ¿Cómo se difunde? ¿Se cita este tipo de literatura? ¿Las tesis doctorales en acceso abierto tienen más visibilidad que la tesis que no son de acceso libre? ¿Las tesis en acceso abierto son más citadas que las tesis que no son de acceso abierto?	Ecuaciones	Nº Resultados	Resultados	
	allintitle: difusión OR diseminación OR citación "literatura gris"	2	https://scholar.google.com/scholar?as_vis=1&q=allintitle:%22difusi%C3%B3n+diseminaci%C3%B3n+citaci%C3%B3n	2
	allintitle: "gray literature" OR "grey literature" dissemination OR diffusion OR	39	https://scholar.google.com/scholar?q=allintitle:%E2%80%9Cgray+literature%E2%80%9	14
3. ¿La interoperabilidad de los repositorios institucionales aumenta la visibilidad y el impacto? ¿Hay una relación directa entre la interoperabilidad de los repositorios institucionales y el aumento de su visibilidad e impacto? ¿Los aspectos técnicos de los repositorios tienen algo que ver en la visibilidad de estos?	Ecuaciones	Nº Resultados	Resultados	
	allintitle: Repositories Interoperability OR visibility OR impact	149	https://scholar.google.com/scholar?as_vis=1&q=allintitle:+Repositories+Interoperabi	50
	allintitle: Repository Interoperability OR visibility OR impact	210	https://scholar.google.com/scholar?q=allintitle:+Repository+Interoperability+OR+visibility+OR+impact&hl=es&as_vlo	7
	allintitle: Repositorios Interoperabilidad OR visibilidad OR impacto	33	https://scholar.google.com/scholar?q=allintitle:%3A+Repositorios+Interoperabilidad+OR+visibilidad+OR+impacto&btnG=&hl=es	5
4. ¿Las políticas de mandato de las instituciones están relacionadas con la visibilidad y el impacto de las mismas	Ecuaciones	Nº Resultados	Resultados	
	allintitle: "open access" mandate OR politics	47	https://scholar.google.com/scholar?as_vis=1&q=allintitle:%22open+access%22+mandate+OR+politics&hl=es&as_sdt=1,5&as_vl	11
	allintitle: acceso abierto políticas OR mandato	25	https://scholar.google.com/scholar?q=allintitle:%3A+acceso+abierto&as_epq=&as_oq=pol%C3%A Dticas+mandato&as_eq=&as_occt=title&as	11
¿Cómo medimos la visibilidad, citación e impacto del acceso abierto? ¿Qué métricas alternativas se utilizan para medir el impacto?	Ecuaciones	Nº Resultados	Resultados	
	almetrics OR informetrics OR scientometrics OR webometrics	4	https://scholar.google.com/scholar?as_vis=1&q=allintitle:%22open+access%22+bibliometrics+OR+almetrics+OR+informetrics+OR	3
nº total de referencias Google Scholar		883		230

Nota: Elaboración propia

La búsqueda en Google Scholar arrojó un total de 883 resultados, que al igual que los resultados de búsqueda en las bases de datos, se gestionaron en *Mendeley* y en *Excel* y se seleccionaron un total de 230 referencias. Una vez hecha la revisión de duplicados el total de referencias relevantes fue de 182 (Anexo 4.Tabla 3.5. Referencias seleccionadas de *Google Scholar*).1.1.4.1.4. Revisar el proceso y evaluar el estudio

La búsqueda de literatura recupera cientos o miles de referencias, a menudo con los resúmenes. Estas deben ser tamizadas para identificar las que son útiles para su posterior revisión.

Se han tamizado los resultados de las búsquedas decidiendo cuáles parecían que cumplían plenamente los criterios de inclusión, y, por tanto, necesitaban un examen más detallado, y cuáles no. Después de excluir las referencias claramente irrelevantes, manteniendo una nota detallada del número de referencias incluidas y excluidas en cada etapa (ver Figura 4), todavía había posibilidades de dejar muchas. Algunas se han podido excluir con confianza después de un examen más profundo del resumen, de otras ha sido necesario obtener el texto completo y se han examinado para determinar si cumplen con los criterios de inclusión y exclusión de la revisión.

Los primeros resultados se cribaron fundamentalmente de acuerdo a dos criterios de inclusión:

1. La información bibliográfica disponible en las bases de datos se utilizó para cribar los resultados por relevancia.
2. Los trabajos a texto completo se obtuvieron por los títulos que parecían más relevantes. Se extrajeron los detalles de cada trabajo y se evaluaron críticamente por la doctoranda y los directores de tesis para garantizar la coherencia. Los trabajos incluidos se evaluaron críticamente utilizando el marco propuesto en este estudio. La figura 4 muestra el diagrama de flujo que describe este proceso.

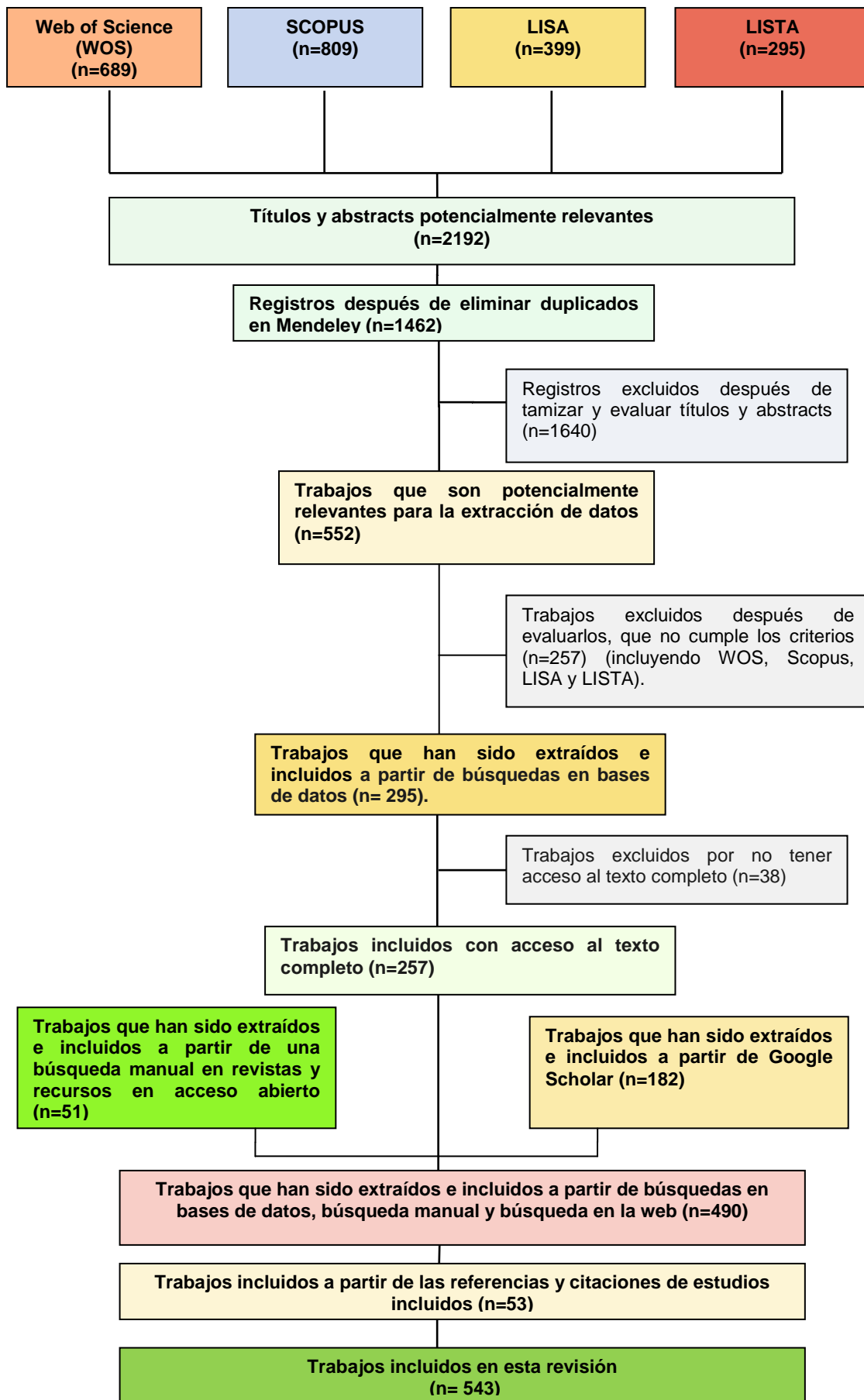


Figura 4. Diagrama de flujo de la selección de estudios para la inclusión en la revisión sistemática. Fuente: Elaboración propia

I.1.4.1.5. Extraer los datos

Una vez completado el cribado, es necesario leer y revisar la literatura que cumple con todos los criterios de inclusión.

En esta etapa de extracción de datos, los elementos específicos que deben ser recogidos varían para cada trabajo y es necesario guiarse por las preguntas y objetivos de investigación. La Tabla 28 muestra el formulario utilizado para la extracción de datos en cada una de las preguntas de investigación efectuadas.

Tabla 28

Formulario de extracción de datos

DATOS EXTRAIDOS	NOTAS
Título del estudio	Identificar el trabajo como un estudio sobre la pregunta de investigación planteada
Resumen	Proporciona un sumario estructurado o un resumen no estructurado como un párrafo. Inclusión de suficientes datos para evaluar la validez de las conclusiones
Autor	
Publicación	Año Fuente Tipo: revista, capítulo de libro, conferencia, proceeding, informe, comunicación, tesis, bibliografía, revisión
Datos sobre acceso al trabajo: Acceso abierto, suscripción.	Si no tenemos acceso al trabajo, se desestimará su adquisición
Pertinencia	Se indicará si el trabajo es pertinente o no
Disponibilidad/Acceso	Se indicará si se tiene acceso al texto completo y en qué forma

Nota: Elaboración propia

Después de este nivel de análisis detallado, se podrán excluir algunos estudios considerados como irrelevantes (Phelps & Campbell, 2012). En el diagrama de flujo (Figura 4) se observa que en este punto se excluyeron 1640 registros extraídos de las bases de datos después de tamizar y evaluar títulos y resúmenes, resultando potencialmente relevantes 552 trabajos para la extracción de datos.

Se ha tratado de sintetizar los trabajos, y evaluar la heterogeneidad entre los hallazgos de los mismos. Las revisiones sistemáticas adoptan un enfoque formal y sistemático para extraer

información relevante de los estudios primarios; lo que a menudo implica el desarrollo de un formulario de extracción de datos, que el revisor completa de todos los estudios en la revisión. Este resume la población, los detalles de la intervención (si los hay), los resultados de interés, y la información metodológica o de otro tipo. Este método tiene por objeto garantizar la coherencia y la objetividad. La extracción de datos también incluye la elaboración de un cuadro detallado que describa todos los estudios que se revisa en detalle (no cada estudio que se encuentra en la revisión sino solo aquellos estudios que cumplan con todos los criterios de inclusión). Después de la evaluación de los 552 trabajos seleccionados, se incluyeron 295 trabajos para ser analizados en detalle a través de la lectura del texto completo. Debido a que no se tenía acceso al texto completo en 38 trabajos, la selección se redujo a 257 trabajos que eran accesibles bien por suscripción o bien por acceso abierto. Por otro lado, se extrajeron datos de 51 trabajos incluidos en la revisión mediante búsqueda manual en revistas y recursos en acceso abierto y por otro lado se extrajeron datos de 182 trabajos incluidos a partir de las búsquedas de las preguntas de investigación realizadas en *Google Scholar*. También se incluyeron 53 trabajos a partir de referencias y citas de los trabajos incluidos. En total la revisión comprende un total de 543 trabajos resultantes del flujo de la selección de la literatura expresado en la Figura 4.

Se usó *Mendeley* para revisar si las referencias importadas cumplían con los criterios de inclusión; organizamos en carpetas las referencias importadas de las distintas búsquedas realizadas en las distintas fuentes (bases de datos, repositorios, recolectores, *Google*, etc.). De los trabajos seleccionados se extrajeron las características de cada trabajo como se ve en la Tabla 28 y se codificaron en *Mendeley*. Las referencias se exportaron desde este gestor con formato de cita *APA* a hojas de cálculo Excel donde se gestionó la duplicidad, se ordenaron los datos de publicación, y se clasificaron los trabajos por pertinencia, tipo de publicación y forma de acceso al texto completo (véase el Anexo 4, Tabla 3.1. *Completo – Criterios de selección*).

I.1.4.1.6. Sintetizar, analizar y presentar los datos

La síntesis de datos implica recopilar y resumir los resultados de los estudios primarios incluidos.

La síntesis puede ser presentada como un sumario o resumen narrativo.

Los estudios incluidos deben integrarse, teniendo en cuenta las variaciones en la población, la intervención (si la hay), el contexto y el entorno, el diseño del estudio, los resultados y el grado en que se ven afectados por el sesgo. Esta integración se puede hacer estadísticamente (meta-análisis) y / o narrativamente, describiendo de manera sistemática la presentación de informes, tabulando e integrando los resultados de los estudios.

En la primera etapa, la síntesis de los datos implica mapear los componentes principales de la literatura, es decir, la tabulación de los resultados para identificar cuántos estudios cumplieron los criterios de inclusión, dónde y cuándo se llevaron a cabo, quiénes fueron sus autores, etc. La siguiente etapa consiste en identificar los temas importantes a través de las referencias. La etapa final consiste en examinar la literatura con más profundidad y explorar la calidad de los estudios de aplicación (sin incluir los estudios introductorios que es la primera categoría), así como poner de relieve las deficiencias en la literatura. Estas tres etapas forman un marco útil para la realización de las revisiones sistemáticas.

En cuanto a la valoración crítica, todos los estudios en la revisión que cumpla con los criterios de inclusión deben evaluarse con respecto a su solidez metodológica. Este proceso ayuda a identificar los sesgos importantes. También ayuda al lector a interpretar los datos. Los resultados de la evaluación crítica se utilizan al sintetizar los resultados de los estudios primarios. La valoración crítica de la literatura seleccionada se irá realizando al ir escribiendo esta memoria ya que constituye la base bibliográfica del mismo.

Algunos indicadores que ayudarán a evaluar la calidad de nuestra revisión sistemática son:

1. ¿Están descritos y son adecuados los criterios de inclusión y exclusión?
2. ¿Es probable que la búsqueda de la literatura recoja todos los estudios relevantes?
3. ¿Han evaluado los revisores la calidad de los trabajos incluidos?
4. ¿Los revisores han tenido en cuenta la calidad del estudio al resumir sus resultados?
5. Si hay un resumen estadístico (meta-análisis), ¿es apropiado? Es decir son los estudios lo suficientemente similares para ser sintetizados estadísticamente?
6. ¿Se evaluó la heterogeneidad del estudio?
7. ¿Fueron respaldadas las conclusiones de los revisores con los resultados de los estudios revisados?

I.1.4.2. Resultados

Como se ha podido comprobar a través de los resultados de la revisión sistemática la literatura sobre el tema del acceso abierto en general es muy abundante, así como los estudios sobre visibilidad e impacto de la investigación en acceso abierto. En cambio la literatura sobre la relación entre literatura gris científica, acceso abierto e impacto es bastante escasa. En la Figura 5 de *Google Trends*¹, se puede ver que el término “*Open Access Impact*” aparece en el año 2011 y su interés se mantiene a lo largo de estos últimos años.

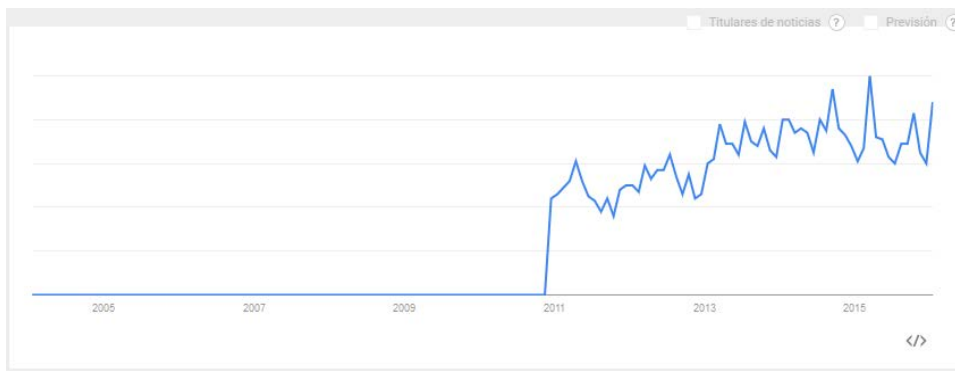


Figura 5. “Open Access Impact”. Fuente: Google Trends

¹ Google Trends, <https://www.google.es/trends/explore#q=open%20access%20impact>

Se han hallado 6 revisiones de la literatura sobre algunos de los temas propuestos en las preguntas de investigación sobre Acceso Abierto que sirvieron para identificar literatura sobre las cuestiones planteadas (ver en el *Apéndice, Tabla VI. Revisiones y bibliografías*). Además de estas revisiones, y como resultado de la revisión del ámbito de los temas genéricos relacionados con el Acceso Abierto, repositorios y preservación digital se localizaron 10 Bibliografías de Charles W. Bailey, Jr., muy útiles para la localización de literatura pertinente en estos temas (*Anexo 4, Tabla 3.6. Revisiones y bibliografías*). En la Tabla 29 se relacionan las revisiones y bibliografías.

Tabla 29

Revisiones y bibliografías

REVISIONES

1. Bar-Ilan, J. (2008). Informetrics at the beginning of the 21st century - A review. *Journal of Informetrics*, 2(1), 1–52. doi:10.1016/j.joi.2007.11.001
2. Bhat, M. H. (2010). Open Access Repositories: A Review. *Library Philosophy and Practice*, NP
3. Craig, I., Plume, A., Mcveigh, M., Pringle, J., & Amin, M. (2007). Do open access articles have greater citation impact? A critical review of the literature. *Journal of Informetrics*, 1(3), 239–248. doi:10.1016/j.joi.2007.04.001
4. Davis, P. M., & Walters, W. H. (2011). The impact of free access to the scientific literature: a review of recent research. *Journal of the Medical Library Association : JMLA*, 99(3), 208–217. doi:10.3163/1536-5050.99.3.008
5. Galina Russell, I. (2011). La visibilidad de los recursos académicos: Una revisión crítica del papel de los repositorios institucionales y el acceso abierto. *Investigacion Bibliotecologica*, 25(53), 159–181
6. Xia, J., Gilchrist, S. B., Smith, N. X. P., Kingery, J. A., Radecki, J. R., Wilhelm, M. L., ... Mahn, A. J. (2012). A Review of Open Access Self-Archiving Mandate Policies. *Portal-Libraries And The Academy*, 12(1), 85–102

BIBLIOGRAFÍAS DE CHARLES W. BAILEY, Jr.

1. [Altmetrics Bibliography \(2013\)](#)
2. [Digital Curation and Preservation Bibliography \(2010-2011\)](#).
3. *Digital Curation Resource Guide* (2012). Esta guía presenta websites seleccionados y documentos sobre curación digital. Complementa a *Digital Curation Bibliography: Preservation and Stewardship of Scholarly Works*
4. [E-science and Academic Libraries Bibliography \(2011\)](#)
5. [Electronic Theses and Dissertations Bibliography \(2005-2012\)](#)
6. [Institutional Repository Bibliography \(2009-2011\)](#)
7. [Open Access Journals Bibliography \(2010\)](#)
8. [Open Access Webliography, con Adrian K. Ho \(2005\)](#)
9. [Research Data Curation Bibliography \(2012-2015\)](#)
10. [Scholarly Electronic Publishing Bibliography \(1996-2011\)](#)

Nota: Elaboración propia

En el *Diagrama de flujo de la selección de estudios para la inclusión en la revisión sistemática* (Figura 4) se presenta el número de referencias identificadas en cada estadio del proceso de revisión de la literatura que se ha llevado a cabo. En total, se identificaron 2.192 referencias de las bases de datos WOS, SCOPUS, LISTA y LISA (sin incluir OA), que se redujeron posteriormente a 552 referencias (*Anexo 4, Tabla 3.2. Completo BBDD*) después de un examen detallado de los títulos y los resúmenes y de la eliminación de duplicados. En la Figura 6 se muestra la tipología de los 552 trabajos hallados.



Figura 6. Tipología de la literatura obtenida en la revisión. Fuente: Elaboración propia

Solo 257 referencias cumplieron con los criterios finales para su inclusión en la revisión (véase Anexo 4, Tabla 3.3. Referencias seleccionadas de las bases de datos WOS, SCOPUS, LISTA y LISA). La mayoría (91%, 233 de 257) de las referencias son artículos de revistas procedentes principalmente de 28 revistas pertenecientes al ámbito de la información y la documentación; las revistas de las que se han obtenido más de cinco referencias se muestran en la Figura 7.

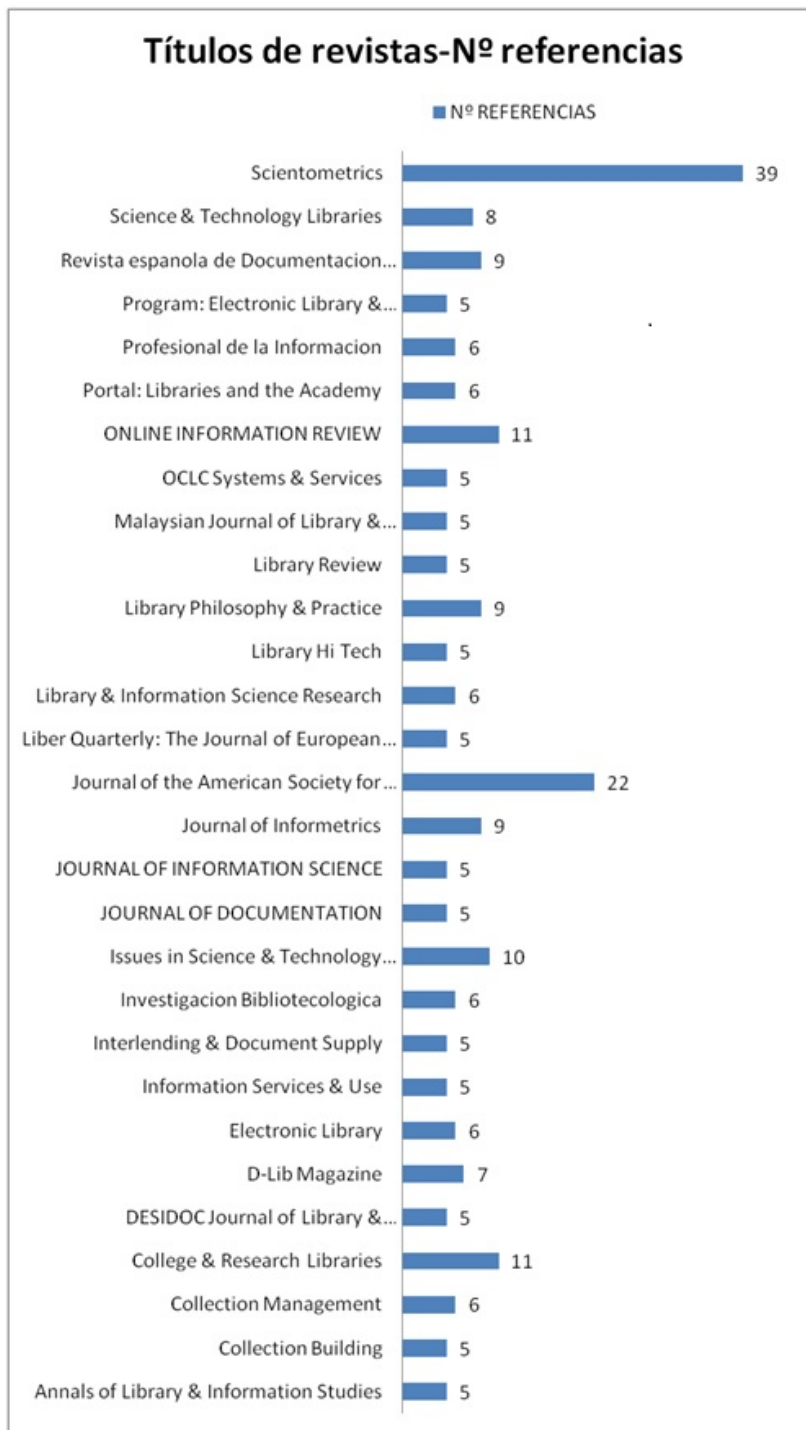


Figura 7. Títulos de revistas más frecuentes y nº de referencias. Fuente: Elaboración propia

Las conferencias ascienden a 17, los capítulos de libros a 3 y a 4 los trabajos clasificados como otro tipo de documentación (carta a editor, noticia de blog, etc.). En la Figura 8 se muestra la relación porcentual de los tipos de trabajos seleccionados.

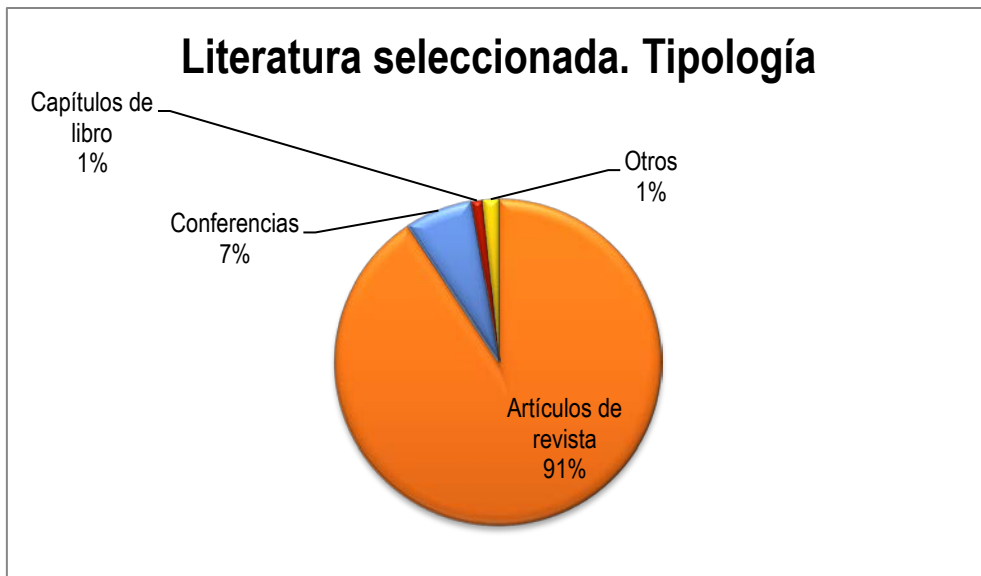


Figura 8. Tipología de la literatura seleccionada. Fuente: Elaboración propia

Las publicaciones están repartidas en un período de tiempo que abarca 1998 a 2015.

Además, 104 referencias se identificaron por métodos complementarios (51 referencias obtenidas de búsquedas manuales y 53 referencias obtenidas a partir de referencias y citas de trabajos incluidos). Para *Google Scholar* y los recursos de acceso abierto, el proceso de revisión fue similar aunque adecuando la estrategia de búsqueda a cada uno de los recursos, dando lugar a 883 referencias de las que se seleccionaron 182. El Anexo 4 incluye las hojas Excel en las que se detallan las diferentes listas de referencias incluidas en el *data set*.

I.1.4.3. Diseminación de la revisión

Un paso importante en el proceso de la revisión sistemática es la diseminación de los resultados de la revisión. La comunicación es fundamental para difundir, usar y reutilizar los resultados de la revisión sistemática. En nuestro caso la revisión sistemática realizada ha tenido desde el principio un objetivo fundamental que era hacer una revisión de la literatura existente sobre las cuestiones planteadas en esta tesis con el fin de fundamentar nuestra investigación sobre una sólida y pertinente bibliografía.

La revisión sistemática realizada ha tenido desde el principio un objetivo fundamental que era hacer una revisión de la literatura sobre las cuestiones planteadas en esta tesis doctoral con el fin de fundamentar la investigación de la misma sobre una sólida y pertinente bibliografía.

Como se explicó al comienzo de este capítulo, no tenía sentido hacer una revisión sistemática que ya se hubiera hecho antes, por ello también se incluyó en la revisión la búsqueda de revisiones sistemáticas ya existentes sobre las cuestiones tratadas; sin embargo, se pudo comprobar a través de los resultados de la revisión sistemática, que mientras que la literatura sobre el tema del acceso abierto en general es abundante, no lo es tanto sobre los estudios de visibilidad e impacto de la investigación en acceso abierto y, es muy escasa sobre los estudios de la relación entre literatura gris científica, acceso abierto e impacto, y sobre revisiones sistemáticas en estos temas.

No se ha incluido la fase de valoración crítica, que conlleva toda revisión sistemática, y que podría ser objeto de una segunda parte de este trabajo. El motivo de no incluirla es que la literatura seleccionada se irá leyendo y analizando en el transcurso de la fase de redacción de la tesis por lo que hasta la finalización de la misma no será posible disponer de la valoración de la totalidad de los estudios seleccionados.

No obstante se considera que este trabajo resultará útil para los investigadores del área ya que como se ha ido refiriendo a lo largo del capítulo, en este campo de investigación la revisión sistemática es todavía un método relativamente desconocido y no se encuentran corpus bibliográficos resultantes de una revisión sistemática de la literatura.

A través de la realización de este trabajo de revisión se han logrado dos importantes objetivos: por un lado estudiar y conocer los tipos de revisiones y, por otro lado, ha sido posible la obtención y selección de una abundante y pertinente literatura sobre los temas de interés en la

que basar la tesis doctoral a través de la aplicación del método de la revisión sistemática de la literatura.

La difusión de los resultados de la revisión se realiza, por tanto a lo largo de las páginas de esta tesis doctoral, a través de todos y cada uno de los módulos y capítulos que la conforman.

Capítulo I.2. El conocimiento científico abierto

El progreso de cada profesión, disciplina académica y de la sociedad en general se apoya en la investigación y en el desarrollo. La investigación genera nueva información y conocimiento.

La información tiene su ciclo de vida: concepción, generación, comunicación, evaluación y validación, uso, impacto, y finalmente se convierte en combustible de nuevas ideas. Los resultados de investigación se publican en revistas, actas de congresos, monografías, tesis, informes, y ahora la web también facilita más de un nuevo espacio para su comunicación.

Desde su origen en el siglo XVII, las revistas científicas son los canales más populares e importantes para la difusión de las nuevas ideas y de la investigación. Las revistas se han convertido en órganos inseparables de la erudición y de la comunicación de la investigación, y son una industria enorme y amplia. Su proliferación (con una alta tasa de mortalidad), el alto coste de la producción, la engorrosa distribución, el tiempo de espera de los autores para ser publicados, el aumento de los precios de suscripción, y por último el archivo de los volúmenes atrasados dieron lugar a un grave problema conocido como “la crisis de las revistas”.

Las TIC, en especial Internet y la Web, han resuelto todos estos problemas bajo el nuevo avatar de las revistas electrónicas que se han hecho inmensamente populares por sus características propias y su versatilidad; pero este nuevo escenario también abre una nueva vía de difusión a otro tipo de literatura de investigación como son las tesis doctorales, piezas clave en la investigación científica y por tanto esenciales en la comunicación de la investigación.

A comienzos del siglo XXI surge el movimiento de acceso abierto con la *Budapest Open Access Initiative* (BOAI), <http://www.budapestopenaccessinitiative.org/>. La filosofía del acceso abierto es proporcionar de forma gratuita y libre el acceso a la investigación y sus publicaciones sin restricciones de *copyright*. El movimiento consiguió el apoyo de científicos, educadores, editores, instituciones de investigación, asociaciones profesionales y organizaciones bibliotecarias. Las

otras declaraciones de acceso abierto, Berlín, <http://openaccess.mpg.de/Berlin-Declaration>, y Bethesda, <http://legacy.earlham.edu/~peters/fos/bethesda.htm>, se asentaron así sobre fuertes cimientos. Su filosofía se basa en que la investigación financiada por los contribuyentes debe estar disponible de forma gratuita para los contribuyentes. La investigación al ser un bien público debe estar disponible para todos, independientemente de la capacidad económica que estos tengan para pagarla. El acceso abierto tiene muchas formas de acceso y uso que varían desde una libertad total, el permiso completo para copiar, descargar, imprimir, distribuir, archivar, traducir e incluso cambiar de formato hasta su uso aplicando diferentes restricciones.

El contexto general en el que se integra este movimiento es el denominado “conocimiento libre” (García-Peñalvo, García de Figuerola, & Merlo-Vega, 2010a), que incluye también el *software* libre, los recursos educativos (Ramírez-Montoya & García-Peñalvo, 2015) o la cultura libre en general (wikipedia, música, etc.), siendo el acceso abierto a la ciencia la parte que se ocupa de los contenidos científicos. El acceso abierto reclama el dominio público para la ciencia y la cultura, que permita la difusión y reutilización del conocimiento y, por ende, el progreso científico y cultural rápido (Abadal, 2012).

Como indica Llorens Largo (2011), la filosofía del conocimiento abierto considera el conocimiento como un bien público de libre acceso. Este mundo nuevo digital 2.0 es el mundo de las cinco Cs: competir, colaborar, crear, comunicar y compartir. En este complejo mundo conectado, los jóvenes han aprendido a convivir y, lo que es más importante, a compartir sus vivencias, experiencias y conocimientos con el resto del mundo. Y es precisamente esta idea de compartir conocimiento la que más está siendo favorecida por la Web 2.0. La necesidad de comunicar conocimiento se beneficia de las herramientas que han surgido en Internet: blogs, wikis, rss, redes sociales, etc. En el mundo 2.0 el mejor reconocimiento lo tiene aquel que se ha dado a conocer gracias a sus aportaciones al mundo.

En el marco de la cultura libre general surgen una serie de tecnologías para la colaboración con fines sociales y el consumo colaborativo, que se denominan tecnologías sociales. Según Saorín y Gómez-Hernández (2014) el conocimiento y uso de estas plataformas permite el empoderamiento de individuos y de grupos, para autoorganizarse para conseguir sus objetivos, participar en la formación de la opinión pública y conseguir beneficios prácticos a través del intercambio apoyado en plataformas digitales

También en el mundo académico se está produciendo esta necesidad comunicativa y de colaboración. En Illanas y Llorens Largo (2011) se realiza una exposición sobre los retos Web 2.0 de cara al EEES. Los autores abundan en la idea de que en la actualidad, cuando los nativos digitales invaden las aulas y los docentes buscan una forma de llegar a ellos, compartir experiencias resulta enriquecedor para ambas partes. Las universidades disponen de plataformas y herramientas directamente destinadas a difundir conocimiento académico y a facilitar que se comparta el conocimiento que se genera dentro de la institución.

Al principio, se dudaba de la autenticidad y calidad de las publicaciones en acceso abierto. Los autores consolidados rehuían tanto a contribuir ellos mismos al acceso abierto como a citar la literatura en acceso abierto. Pero el *Committee on Publication Ethics* (COPE, 1997)² a través de su código de conducta formulado en colaboración con DOAJ³ y OASPA⁴ condenaron estas conductas. Definieron buenas prácticas y reunieron principios de transparencia para el control de calidad con el fin de evitar el fraude. Ahora se ha demostrado y se acepta que los contribuyentes al acceso abierto aumentan su visibilidad, la presencia global, la accesibilidad, la colaboración, el impacto, tanto en citas como en aplicaciones, y en última instancia la realimentación con comentarios y reflexiones críticas. Este movimiento tuvo sus raíces en una campaña de promoción sistemática. Cada año, desde 2008, en la última semana de octubre se viene

² COPE, <http://publicationethics.org/>

³ DOAJ, <https://doaj.org/>

⁴ OASPA, <http://oaspa.org/>

celebrando la semana internacional del acceso abierto (OAW)⁵ en todo el mundo. Existen muchas organizaciones que promocionan este movimiento a través de los medios sociales y proporcionan una guía de buena práctica para el resto de organizaciones.

Según el trabajo publicado por la UNESCO en 2015 (Kanjilal & Das, 2015), el acceso abierto a la literatura científica no solo permite que la difusión de nuevas ideas sea más fácil y rápida, sino que el impacto de la investigación pueda ser medido cuantitativamente por diversos métodos bibliométricos, cienciométricos y webométricos como el índice h y el índice i-10 con el fin de medir la productividad científica, su flujo, su velocidad y finalmente su influencia concreta sobre individuos y sobre el progreso de una disciplina. El movimiento de acceso abierto está ganando terreno día a día, gracias a la tecnología, y a los esfuerzos para controlar la calidad y medir el impacto de la productividad y de la investigación, convirtiéndose en un modelo eficaz y eficiente para la comunicación científica.

I.2.1. La comunicación científica

El término “comunicación científica” describe el proceso de compartir y publicar trabajos y resultados de investigación (Borgman & Furner, 2005). Mediante la comunicación científica es posible poner la investigación a disposición de una comunidad científica más amplia (Halliday, 2001).

De acuerdo a García-Peñalvo (2016b) la escritura de artículos es la forma más común de comunicación científica. Cuanto mejor sea la capacidad de transmitir y comunicar de un investigador mayores serán los beneficios para la carrera personal del propio investigador, para la investigación en sí y para las instituciones que financian y reciben los resultados de la investigación. El mismo autor afirma que la comunicación científica debe ayudar a disminuir la distancia entre la sociedad y la academia.

⁵ OAW, <http://www.openaccessweek.org/>

García-Peñalvo (2016b) también distingue entre diferentes formas de comunicación científica: diseminación, difusión y divulgación científica. La diseminación científica se refiere a la transmisión, por parte de los investigadores, de informaciones científicas y tecnológicas para sus pares o especialistas en el mismo sector de la Ciencia utilizando un lenguaje específico. La difusión científica se refiere a la misión de los investigadores de transmitir al público los conocimientos de su disciplina, donde el público incluye a profesionales de otras áreas. La divulgación científica se refiere a la transmisión al gran público, empleando un lenguaje accesible, informaciones científicas y tecnológicas.

Para Graham (2000) el proceso de comunicación científica se divide en tres etapas principales: la comunicación en redes de trabajo informales, como por ejemplo los medios de comunicación social, la diseminación pública inicial a través de conferencias y *preprints*, y la publicación formal de la investigación en revistas científicas de investigación. En este proceso hay varios actores involucrados: los autores, los editores, las bibliotecas, los investigadores, las instituciones de educación superior, y las agencias de financiación.

Según Veiga de Cabo y Martín-Rodero (2011) la ciencia se basa en la aportación de los hallazgos y descubrimientos precedentes y en su integración en un conjunto de conocimientos compartidos por los científicos y que constituye su paradigma. Pero además la ciencia se basa en la comunicación y difusión de los resultados. No se puede concebir la ciencia sin que haya comunicación y transmisión de conocimientos. La comunicación científica es la forma por la que los investigadores, científicos y docentes comparten su trabajo con el resto de la comunidad académica, desempeñando un papel fundamental en la diseminación del conocimiento resultante de la investigación continua y de la experimentación. La necesidad de hacer público el conocimiento científico se formaliza con la Revolución Científica en el siglo XVII y se mantiene hasta el momento actual, debido, entre otros factores, a la progresiva institucionalización y profesionalización de la ciencia, especialmente desde mediados del siglo XIX.

La comunicación de la información comienza con la civilización cuando las personas grababan sus experiencias en inscripciones y más tarde en manuscritos. La invención de la imprenta por Gutenberg en la mitad del siglo XV supuso la revolución de la publicación en forma de libro impreso y constituyó un avance en la difusión del conocimiento científico a gran escala así como la condición necesaria para el nacimiento de la ciencia moderna (Kanjilal & Das, 2015).

Desde la aparición de las dos primeras revistas científicas a mediados del siglo XVII, *Journal des Sçavans* y *Philosophical Transactions of the Royal Society⁶ of London*, los investigadores y científicos se han esforzado por publicar y difundir los resultados de su trabajo de investigación a través de revistas científicas, convirtiéndose en una plataforma para compartir sus hallazgos y en un registro público de comunicación científica. Según Martín-Rodero (2014) en ese tiempo los científicos sintieron la necesidad de publicar sus descubrimientos y hallazgos teniendo como motivos principales, por un lado el deseo de reconocimiento y por otro la preservación de la autoría.

La comunicación científica tiene una larga historia desde los comienzos de la *Royal Society* de Londres, reconocida como el primer editor del mundo científico al publicar en 1665 la primera revista científica con revisión por pares. La sociedad fue considerada como un lugar de investigación y discusión. La publicación de la primera revista habilitó las prácticas para continuar fuera de la Royal Society (Syfret, 1948). Esta sociedad continúa representando un importante papel en la comunidad científica y cuenta con destacadas figuras asociadas a ella que han sido reconocidas por sus contribuciones a la ciencia.

Las revistas científicas se erigieron así en el vehículo perfecto para la comunicación de los nuevos avances científicos, quedando asegurada la calidad de lo publicado a través de un sistema de revisión por pares independiente. El documento científico por excelencia es el artículo

⁶ Philosophical Transactions, <http://rstl.royalsocietypublishing.org/>

publicado en una revista científica cuyo objetivo prioritario es comunicar y difundir los resultados de la investigación. Sin embargo, junto a los llamados canales formales de comunicación científica como revistas científicas, libros y actas de congresos existe la literatura gris científica, entre la que se encuentran las tesis doctorales, objeto de estudio de esta tesis.

Para la mayoría de los científicos del mundo el estándar por excelencia de comunicación científica es el artículo revisado por pares. La publicación de artículos en revistas revisadas por pares es el principal indicador del estatus profesional para los investigadores; además de cumplir con otros requisitos, tales como el reconocimiento de autor, el control de calidad, el registro histórico y el archivo para el progreso de la ciencia. Este cúmulo de ventajas dio lugar al crecimiento sin precedentes de las revistas *peer-review*.

A partir del crecimiento exponencial de la literatura científica desde la Segunda Guerra Mundial, se hizo sentir la necesidad de crear nuevas vías de organización, almacenamiento y acceso a la enorme cantidad de información. La invención por Vannevar Bush del sistema basado en el microformato, *Memex*, para almacenar y recuperar la información usando una serie de enlaces navegables es considerado como el precursor del hipertexto moderno basado en sistemas electrónicos (Kanjilal & Das, 2015).

La expansión de la actividad I+D ha sido una de las causas del enorme crecimiento de publicaciones en los últimos años. Desde entonces las instituciones académicas y de investigación de todo el mundo han estado luchando con los problemas y cuestiones relacionados con el proceso de comunicación. Los principales problemas son los relacionados con la gestión de la gran explosión de la información, lo que produce un aumento de los costes de la edición, retrasos en la publicación e ineficacia en la distribución. Por otra parte, las bibliotecas se enfrentan al problema del aumento del precio de las revistas, al espacio físico limitado para el almacenamiento y a la escasez de recursos. La publicación electrónica y el

procesamiento digital de la información, así como su almacenamiento y recuperación ha tenido un gran impacto en el proceso de comunicación tanto desde el punto de vista de la edición como de la diseminación. La aparición de las revistas electrónicas en los años 80 y el desarrollo de la Web en los 90 revolucionaron el panorama de la comunicación científica.

El desarrollo de Internet ha tenido grandes implicaciones en la diseminación y publicación científica (Walsh, Kucker, Maloney, & Gabbay, 2000). El formato electrónico ha influido enormemente en la publicación científica, sobre todo desde el punto de vista de la disponibilidad y la accesibilidad (Björk, 2007). Aunque la distribución de la información científica haya conservado parte de sus estructuras tradicionales, las vías de la comunicación científica y de la difusión de la investigación se han visto afectadas de manera sustancial a través de una mayor comodidad, fiabilidad y bajo coste de producción de la información (R. Campbell & Meadows, 2011). Con la llegada de tecnologías facilitadoras y los modelos innovadores de edición científica que están emergiendo se está facilitando la auto-publicación en la que la responsabilidad y la propiedad de la investigación descansan sobre los creadores de la misma. La web social e interactiva conlleva desafíos y ventajas para la comunicación científica, con el establecimiento del nuevo tipo de sociedad en red (Dahlgren, 2005; Katz, Rice, & Aspden, 2001), los investigadores conocen una evolución en la comunicación científica, lo que requiere un mayor conocimiento de toda clase de procesos de comunicación incluso cuando trabajan en solitario (Thorin, 2006).

Según Ho (2016) la comunicación científica es un proceso cíclico en el que el contenido es generado, revisado, diseminado, adquirido, preservado, descubierto, recuperado y asimilado para el avance de la ciencia. La asimilación puede conducir a la generación de nuevo contenido y de esta forma se puede volver a comenzar una nueva iteración del proceso o ciclo de vida.

La comunicación científica hace referencia a la creación, transformación, evaluación (*peer-review*) diseminación y preservación del conocimiento científico. Es el componente vital del proceso de la investigación. El método más común de la comunicación científica hasta hace muy poco tiempo ha sido a través de la redacción de los hallazgos de la investigación en un libro o en un artículo publicado en una revista académica. Pero como se ha explicado anteriormente con la llegada de Internet y otras aplicaciones de las TIC se produce un cambio importante en el proceso de la comunicación científica. Se observa un cambio en los procesos de publicación en donde están los investigadores que utilizan una gran variedad de medios y formatos con el fin de compartir y difundir sus obras. El entorno digital en red ha permitido la creación de plataformas para que los investigadores publiquen allí directamente y estas se están convirtiendo en herramientas esenciales para los estudiosos que realizan investigación, creando redes científicas donde difundir sus ideas y trabajos. Las bibliotecas desempeñan un papel importante en este proceso de la comunicación científica. El acercamiento a la comunidad académica es un hecho relevante en la actividad de las bibliotecas para facilitar un cambio positivo hacia el desarrollo del sistema de comunicación científica a través de nuevos modelos de investigación y difusión.

El diagrama del ciclo de vida de la investigación (Figura 9) elaborado por Joint Information Systems Committee (2014) muestra dos círculos interconectados, el primero muestra el ciclo de vida de la investigación y del segundo emana el ciclo de vida de los datos desde la etapa de proceso de investigación.

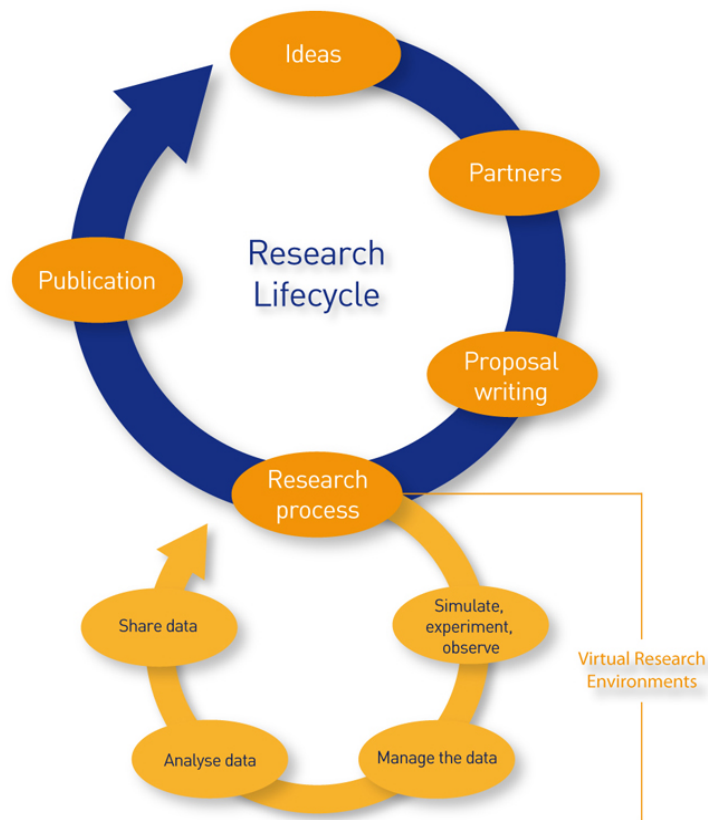


Figura 9. Ciclo de vida de la investigación (JIRCS), <http://www.jisc.ac.uk/whatwedo/campaigns/res3/jisichelp.aspx>

Según la ARL la comunicación científica puede definirse como el sistema a través del cual la investigación y otros escritos científicos son creados, evaluados, diseminados a la comunidad científica y preservados para su futuro uso. El sistema incluye dos medios de comunicación: los canales formales de comunicación, como son la publicación en revistas, y los canales informales, como son los servidores de listas electrónicas (Association of College & Research Libraries, 2003).

Hoy en día las TIC han revolucionado la forma de recoger, almacenar, compartir y comunicar la información. Esto también ha tenido impacto en el proceso de comunicación científica sobre todo desde el punto de vista de la publicación y la difusión (Withey et al., 2011).

El sistema de la publicación científica en la actualidad es una agregación que combina cuatro funciones de la publicación de revistas: registro, certificación (*peer review*), conocimiento (comunicaciones) y archivo en un solo paquete. Internet y otras aplicaciones TIC proporcionan

una oportunidad para desbloquear el sistema de la publicación científica tradicional y procura nuevas formas para cumplir con estas cuatro funciones.

Actualmente una parte sustancial de las publicaciones científicas están controladas por un número limitado de grandes editoriales. Este hecho ha impactado directamente en el proceso de la comunicación científica y amenaza con destruir el propósito para el cual la comunidad académica lo creó. En las últimas dos décadas se ha producido un aumento del número de revistas científicas, una escalada de los precios (Kyrillidou & Young, 2008), y la caída del poder adquisitivo de los países en desarrollo; como consecuencia se produjo una situación de crisis en la que las instituciones académicas y de investigación no podían suscribirse a todas las publicaciones que le interesaban viéndose forzadas a cancelar suscripciones y a recurrir a otros métodos para facilitar el acceso a la comunidad científica. Esta situación allanó el camino para el auge del movimiento por el acceso abierto (Kanjilal & Das, 2015).

1.2.2. El acceso abierto como nuevo modelo de comunicación científica

Entre esos nuevos modelos de comunicación científica que surgen con la irrupción de Internet y las TIC se encuentra el acceso abierto al que Abadal (2012) considera como una revolución, un cambio de paradigma en el funcionamiento de la comunicación científica que no es ni gratis ni libre, ya que para su uso hay que pagar unas licencias, estando la mayor parte de sus contenidos bajo derechos de explotación de las editoriales.

Una de las definiciones más conocidas de acceso abierto es la de Peter Suber: *"Open-access literature is digital, online, free of charge, and free of most copyright and licensing restrictions"* (Suber, 2004). La literatura en acceso abierto (OA) es digital, en línea, gratuita y libre de restricciones debidas a los derechos de explotación y a las licencias de uso. Esto quiere decir que las dos condiciones para que los contenidos científicos se consideren de acceso abierto son que sean gratuitos y, además que estén libres de algunas restricciones de derechos de

explotación. El acceso abierto al conocimiento representa un cambio radical de modelo en el funcionamiento de la comunicación científica.

Hace algo más de una década que surgió el movimiento de acceso abierto al conocimiento cuyos principios se perfilaron en la *Budapest Open Access Initiative* de 2002 (BOAI) y se reafirmaron en 2012 con motivo de la celebración de los diez años de vida de este movimiento⁷. Fue en la Declaración de Budapest donde se definieron cuáles iban a ser los elementos esenciales del acceso abierto: “Una vieja tradición y una nueva tecnología convergen para hacer posible un bien público sin precedentes. La vieja tradición es la voluntad de científicos y académicos de publicar los frutos de sus investigaciones en revistas científicas sin remuneración, solo por el bien de la investigación y la difusión del conocimiento. La nueva tecnología es Internet. El bien público que las dos hacen posible es la distribución digital a todo el mundo de la literatura científica revisada por expertos así como el acceso totalmente libre y sin restricciones a ella para todos los científicos, académicos, profesores, estudiantes y otras personas interesadas”.

La “vieja tradición” se refiere a la comunicación científica y como se ha podido ver anteriormente en este capítulo tiene sus orígenes en la segunda mitad del s. XVII, siendo su función la de contribuir al avance social de la ciencia a la vez que preservar y dar a conocer la autoría de estos progresos científicos. Por otro lado, la “nueva tecnología” que es Internet, ha hecho posible grandes cambios en el acceso a la información, a la cultura, al ocio y al entretenimiento, ofreciéndonos la oportunidad de construir una representación global e interactiva del conocimiento humano, incluyendo el patrimonio cultural, y una perspectiva de acceso mundial (López, 2013).

⁷ “Diez años desde la Budapest Open Access Initiative: hacia lo abierto por defecto”
<http://www.soros.org/openaccess/boai-10-translations/spanish>

Los efectos generados tanto por internet y como por la digitalización de contenidos unidos al sistema de comunicación científica han sido fundamentales para facilitar el desarrollo del acceso abierto, haciendo posible que los contenidos científicos digitalizados sean difundidos inmediatamente y a bajo coste a través de internet (Abadal, 2012). El acceso abierto se convirtió en una realidad cuando un gran número de instituciones unieron sus fuerzas para promover la libre diseminación de la producción científica y empujaron a las administraciones públicas a crear repositorios digitales que pudieran ser consultados libremente (García-Peñalvo et al., 2010).

Hace veinte años, la idea del acceso abierto correspondía sobre todo al concepto de la comunicación directa, sobre el modelo de la plataforma arXiv considerada como el primer archivo abierto. Inmediatamente aparecieron las primeras revistas en acceso abierto tales como *Nucleic Acids Research*, *New Journal of Physics* y *Atmospheric Chemistry and Physics*. Stevan Harnad definía estas dos formas como *green road* y *gold road* (Stevan Harnad et al., 2008). La vía verde se refiere al autoarchivo de las publicaciones en los repositorios mientras que la vía dorada corresponde a la publicación de artículos en revistas de acceso abierto (Baruch, 2007; Guédon, 2008b; Schöpfel & Prost, 2013).

En pocos años, la vía dorada se ha transformado en un nuevo mercado donde cohabitan iniciativas locales de nuevos modelos económicos y la oferta comercial de los editores, que hacen que los autores (o instituciones) tengan que pagar por la libre difusión de sus publicaciones. La vía verde se ha desarrollado también en varias direcciones. Armbruster y Romary (2010) distinguen cuatro tipos diferentes de repositorios: los temáticos, los nacionales, los institucionales, y otros sitios de carácter científico que no pertenece a ninguna de estas categorías. De todas formas a veces es difícil hacer la distinción entre estas cuatro categorías, también en relación con otros sitios científicos (bibliotecas digitales, portales, etc.).

En cuanto a las dos rutas del acceso abierto, Pinfield (2015) argumenta que la dualidad verde-dorada es una herramienta bien establecida del discurso del acceso abierto. Es la base de gran parte de la literatura sobre el acceso abierto y del debate de su política. Sin embargo, una característica clave del paisaje actual del acceso abierto es la continua incertidumbre y desacuerdo acerca de la relación entre las rutas verde y dorada. Suber (2012) argumentó que las dos son “complementarias y sinérgicas” tanto a corto como a largo plazo. Guédon (2008a) también argumentó un caso similar. Pinfield (2009) ha presentado muchos modelos de la relación sinérgica entre repositorios y revistas basada en un análisis de desarrollo de servicios de acceso abierto. En la práctica, muchas instituciones han trabajado para apoyar tanto la vía verde como la dorada, creando y gestionando repositorios al mismo tiempo que diseñaban procesos para facilitar el pago de los gastos por publicar (APC, *Article Processing Charges*). Sin embargo, está implícito en muchos de los actuales debates entre defensores del acceso abierto que las dos rutas son rivales. Este debate quedó claro tras la publicación del Informe Finch en el Reino Unido en 2012⁸ (Mabe & Price, 2012). Stevan Harnad (2012) tenaz defensor del acceso abierto verde ha criticado cualquier movimiento hacia la ruta dorada. Aunque muchos defensores de la ruta dorada han sido más cautelosos al expresar sus opiniones, continúan manifestando su escepticismo sobre la validez de los repositorios. El debate, por tanto, continúa, y, aunque los términos del mismo parecen estar de acuerdo ya que hay un entendimiento implícito de que verde y dorada son las dos rutas principales posibles hacia el acceso abierto, no hay acuerdo acerca de cómo podría y debería desarrollarse el acceso abierto. Esta tensión subyace en gran parte del discurso actual y parece que continuará durante largo tiempo.

Según Brody (2006) la investigación científica es un terreno competitivo, en el que los investigadores tendrán que demostrar que su trabajo tiene impacto en su campo de investigación con el fin de asegurar la financiación, la promoción y la permanencia. Para maximizar el impacto,

⁸ Finch Report, <http://www.researchinfonet.org/publish/finch/>

los investigadores se comprometen a priorizar la alta investigación, primero persiguen obtener resultados, y después intentan publicar en revistas de alto impacto. Internet representa una nueva oportunidad para el autor académico que busca el mayor impacto posible: ahora podrá hacer su trabajo accesible al instante en la Web a través del autoarchivo en repositorios. Este crecimiento de la literatura en acceso abierto (junto con nuevos modelos de publicación que hacen que las revistas estén disponibles libremente) maximiza el impacto de la investigación al maximizar el número de personas que pueden leer, y al estar disponible mucho antes. El acceso abierto también ofrece un nuevo objeto de estudio para la investigación bibliométrica.

Pinfield (2015) plantea que uno de los desafíos relacionados con el acceso abierto es el desarrollo de pruebas o evidencias que puedan fundamentar los debates actuales acerca de cómo será la comunicación científica en el futuro, es decir si el acceso abierto puede ser el modelo por excelencia para la comunicación científica. Mientras que las evidencias todavía necesitan ser reforzadas en muchos aspectos, el trabajo realizado hasta la fecha ya muestra la creciente importancia del acceso abierto en el entorno de la comunicación científica. Los estudios llevados a cabo sobre el crecimiento de los repositorios (número y contenido), utilizando diversas fuentes, que incluyen, por ejemplo, OpenDOAR, <http://opendoar.org/>, y BASE, <https://www.base-search.net/>, son indicativos de la importancia del acceso abierto (Björk, Laakso, Welling, & Paetau, 2014; Pinfield et al., 2014).

Schöpfel y Prost (2013) exponen en uno de sus trabajos que desde la Declaración de Berlín de 2003, que definió el acceso abierto como “una fuente universal del conocimiento humano y del patrimonio cultural que contaba con la aprobación de la comunidad científica”, hasta 2013 el número de repositorios ha sobrepasado los 2.000, se cuenta con más de 8.000 revistas en acceso abierto, y es imposible decir cuántos artículos, comunicaciones, tesis doctorales, etc. estarán en adelante accesibles libremente. Una parte de la comunicación científica se ha convertido en “bien público” con una difusión mucho mayor que antes (Willinsky, 2005). Los

autores comentan que a pesar del éxito innegable de este movimiento en términos de cifras, sin embargo no se trata de un tema coherente, uniforme y coordinado, al contrario, la diversidad es la regla tanto en lo que respecta a los logros técnicos como a los modelos económicos (Schöpfel & Prost, 2013; Alma Swan & Awre, 2006).

Los datos sobre acceso abierto ofrecidos por Schöpfel y Prost (2013) se pueden ir actualizando periódicamente con los datos sobre acceso abierto publicados por Heather Morrison en el blog *The Imaginary Journal of Poetic Economics* (Morrison, n.d.), cuya última actualización a la que se ha tenido acceso es la de 30 de junio de 2016. En cuanto a datos de artículos y revistas OA, a pesar del crecimiento negativo de conjunto que emana de un proyecto de limpieza de gran mano dura, el número de artículos recuperables desde DOAJ creció un 16% en 2016 con relación al año 2015, situándose en 300.000 artículos de un total de 2,1 millones. El 9 de mayo de 2016, DOAJ retiró más de 3.000 revistas que no habían rellenado el nuevo formulario de solicitud. Desde esa fecha, DOAJ ha ido añadiendo títulos que cumplían con los requerimientos a razón de 4,5 títulos por día, ascendiendo el número de revistas a más de 9.000 (según datos de agosto 2016 del portal de DOAJ). Las cifras del *Bielefeld Academic Search Engine* (BASE), <https://www.base-search.net/>, demuestran el crecimiento global en la mayoría de los repositorios de acceso abierto, añadiendo más de 18 millones de documentos en relación al pasado año hasta alcanzar un total de 93 millones de documentos, y añadiendo más de 670 proveedores de contenido con un total de 4.000 proveedores de contenido. Tanto el crecimiento de documentos como el crecimiento de proveedores de contenido en BASE refleja un crecimiento superior al 24% en el año 2016. *The Internet Archive*, <https://archive.org/index.php>, alcanza en junio de 2016 más de 10 millones de textos. Lo más importante es que más de 10 millones de los textos puestos a disposición por *Internet Archive* y *Open Library*, <https://openlibrary.org/>, son totalmente accesibles y de dominio público.

En los últimos cinco años se han publicado otros estudios que tratan sobre la proporción de la literatura publicada que está en acceso abierto. (Björk et al., 2010) estiman que el 20% de los trabajos publicados en 2008 estaban en acceso abierto mientras que Gargouri et al. (2012) estiman el 22%. Un estudio realizado por Elsevier (2013) identifica niveles de captación global en varios modelos de acceso abierto: totalmente acceso abierto dorado, 5,5%; híbrido 0,5%; acceso abierto verde del manuscrito final aceptado 5%, etc. Pero señalaron que estas tasas de absorción no se pueden resumir de manera significativa a través de estos modelos debido a los diferentes métodos de medición y períodos, y también debido a la posible duplicación a través de diferentes modelos. También en 2013, otro estudio financiado por la Comisión Europea y llevado a cabo por Archambault et al. (2013) sugiere que el acceso abierto está alcanzando un punto de inflexión basado en la estimación de que 48% de la literatura publicada en 2008 estaba disponible en acceso abierto en 2012. El estudio recoge 22 áreas disciplinares y muestra una considerable variación en cada una de ellas, pero en algunas disciplinas la mayor parte de los artículos están ya en acceso abierto incluyendo ciencia y tecnología (64%), investigación biomédica (61%), biología (57%) y matemáticas y estadística (56%). Por otra parte, Chen (2014) ha realizado un estudio cuyos resultados indican unos niveles similares de acceso abierto en la literatura. Todo parece indicar que hay un crecimiento continuo en el acceso abierto desde una variedad de perspectivas hasta un punto en que se puede decir que el acceso abierto está entrando en la corriente principal de la comunicación científica.

En cuanto a la literatura tratada en este trabajo de investigación (las tesis doctorales), los datos también apuntan a un creciente interés por difundir en acceso abierto este tipo de literatura gris. *Networked digital Library of Theses and Dissertations* (NDLTD) contiene ya 4.434.714 tesis electrónicas, *Open Access Theses and Dissertations* (OATD) indexa más de 3.100.000 de tesis, DART Europe da acceso abierto a 680.782 tesis de 592 Universidades en 28 países europeos y

Tesis Doctorales en Red (TDR) contiene más de 23.300 tesis depositadas a texto completo y más de 76.300 tesis consultables.

I.2.3. Las ventajas del acceso abierto al conocimiento

Según Suber (2012) el acceso abierto beneficia a todos, por las mismas razones que la investigación en sí beneficia también a todo el mundo. El acceso abierto cumple esta función, haciendo que la investigación y los resultados de la misma estén disponibles y sean útiles. Beneficia a los investigadores como lectores, ayudándoles a buscar y recuperar la información que necesitan, y beneficia a los investigadores como autores, porque les permite llegar a los lectores que pueden citar y basar su trabajo en otros previos. El acceso abierto beneficia también a sectores que no se dedican específicamente a la investigación, acelerando los logros y beneficios derivados de la investigación.

Las declaraciones *Open Access* (OA) Budapest-Bethesda-Berlin (BBB) y la literatura asociada identificaron los principales beneficios de la publicación científica en acceso abierto. El principal beneficio es proporcionar la eliminación de barreras de acceso, ya que no hay costes asociados con suscripciones o el acceso en línea. No hay ninguna restricción de acceso en modo OA que conduzca a la reducción de los conocimientos. Muchos de los beneficios de la publicación en acceso abierto están relacionados con la naturaleza global de internet. Internet ofrece cada contenido *online* para un público general o para un público especializado, dependiendo de la naturaleza del contenido. Estos contenidos libres son consultados y apreciados por las comunidades globales. Por tanto la publicación en acceso abierto brinda todo el potencial de la investigación difundida. Para un autor esto significa el aumento de visibilidad global, el aumento de la accesibilidad tanto en países desarrollados como en países en desarrollo, incrementa las posibilidades de ser citado e incrementa las métricas a nivel de artículo (métricas alternativas). Otros investigadores pueden encontrar colaboraciones de investigación relacionadas con sus

publicaciones para posibles colaboraciones internacionales a nivel institucional o individual. Los contenidos abiertos también ofrecen una ventana para recibir opiniones constructivas y comentarios. Estas reflexiones críticas enriquecen la publicación de contenidos en acceso abierto. Tanto lectores como autores pueden ser parte activa en el avance del conocimiento ofrecido por la literatura científica abierta.

Houghton (2009) también hace referencia a que el acceso abierto introduce la investigación nacional en una red de trabajo interoperable de conocimiento global, incrementa el impacto de la investigación nacional, provee nuevos colaboradores de investigación, y elimina el aislamiento profesional. La sociedad en su conjunto se beneficia también de ello ya que el acceso abierto facilita que la investigación sea más eficiente y más eficaz y que se puedan obtener resultados más rápidos y mejores para todos. El acceso abierto fortalece las economías mediante el desarrollo de una base nacional de ciencia fuerte e independiente. Existe una creciente evidencia de que los países también se benefician porque el acceso abierto aumenta el impacto de la investigación en la que invierten el dinero público y, por tanto, hay un mejor retorno de la inversión.

Es notorio que el acceso abierto tiene repercusiones en cualquier campo de la economía, de la industria y de la tecnología, pero es claramente en el campo de la educación y de la producción científica donde este movimiento genera una evolución que afecta a muchos modelos colaterales de negocio. El acceso abierto se ha perfilado como un cambio radical en la diseminación de los resultados científicos y de transferencia en el sector de producción hacia la innovación (García-Peñalvo et al., 2010a).

El acceso abierto, en la actualidad, está ampliamente apoyado por los gobiernos y los organismos de financiación que muestran actitudes cada vez más favorables a la apertura de datos y al acceso abierto a los contenidos. No cabe la menor duda de que el acceso abierto es

una forma de amortizar la inversión en investigación de un país, al poner a disposición de los usuarios, la documentación derivada de la misma, además de ayudar a la inclusión digital de los ciudadanos al ofrecer los contenidos a texto completo a los investigadores, los docentes, los estudiantes, las instituciones y a toda la ciudadanía (Ferrerías-Fernández, 2011a).

Significativo es que al igual que en el anterior 7º Programa Marco, en el programa marco europeo de trabajo para la investigación, *Horizon 2020*, se trate el tema de *Open Access* y *Open Data*, haciendo hincapié en que la investigación y la innovación son vitales para el futuro de las personas, y por ello es importante ofrecer libremente al público las actividades de investigación y de innovación financiadas por *Horizon 2020*⁹.

Y es que como bien apunta Abadal (2012), las ventajas de este nuevo modelo son muchas y se podrían agrupar en tres grandes apartados:

1. Los resultados de investigación y de las publicaciones científicas puestos a disposición en acceso abierto supone una notable mejora en el funcionamiento de la comunicación científica. Aumenta el uso e impacto de los contenidos, se mejora la calidad de la investigación y pueden reducirse notablemente los costes a través de este modelo de publicación.
2. El acceso abierto también genera beneficios directos sobre la sociedad ya que facilita una transferencia directa de conocimiento a la sociedad, rompe las barreras entre países ricos y pobres, y permite visibilizar la inversión pública en investigación.
3. El acceso abierto posibilita la reutilización de la información y de los datos. Los contenidos se ponen a disposición no solo para la consulta sino también para que puedan reutilizarse creando productos y servicios derivados. Para que esta reutilización sea posible es necesario que se hayan liberado las barreras legales y, por otra parte,

⁹ Horizon 2020, http://ec.europa.eu/research/horizon2020/index_en.cfm

que se cumplan determinadas especificaciones técnicas para facilitar la interoperabilidad entre productos y sistemas.

En relación a los beneficios de aumento de citación, según Niyazov y otros (2016) muchos investigadores, comenzando por Lawrence (2001), han encontrado que los artículos en libre acceso tienden a tener más citaciones que los artículos de pago. Esta ventaja de citación se ha observado en una serie de estudios, que abarca una variedad de campos académicos incluyendo la informática (Lawrence, 2001), la física (Harnad & Brody, 2004), y la biología y la química (Eysenbach, 2006).

Pinfield (2015) hace referencia a que se han desarrollado evidencias sobre los efectos del acceso abierto (o sus efectos potenciales). El ejemplo más claro de esto es la ventaja de citación. Una amplia variedad de estudios (la mayoría de ellos llevados a cabo sobre este tema), en diferentes disciplinas y utilizando diferentes metodologías, han encontrado que existe ventaja de citación para los trabajos en acceso abierto, aunque la magnitud de la ventaja varía entre los estudios realizados (Archambault et al., 2013; Atchison & Bull, 2015; Bernius & Hanauske, 2009; Calver & Bradley, 2010; W. Cheng & Ren, 2008; Davis, 2009, 2011; Davis, Lewenstein, Simon, Booth, & Connolly, 2008; J.M. Donovan & Watson, 2011; Evans & Reimer, 2009; Frandsen, 2009; Frisch et al., 2014; Gaulé & Maystre, 2011; Ingwersen & Elleby, 2011; K. & M., 2010; Koler-Povh, Juznic, et al., 2014; Kousha & Abdoli, 2010; Kullman, 2014; Laakso & Björk, 2013; Lansingh & Carter, 2009; S.-K. Lin, 2009; M., C., F., et al., 2008; M., C., & F., 2008; McCabe & Snyder, 2014; Mueller-Langer & Watt, 2014; Piwowar & Vision, 2013; Riera & Aibar, 2013; Sabharwal, Patel, & Johal, 2014; Snijder, 2010; Solomona, Laaksob, & Bjork, 2013; Vanclay, 2013; Wagner, 2010; Xia et al., 2011).

En contraste, otra interesante línea que actualmente está recibiendo la atención de los científicos es el trabajo que se lleva a cabo para analizar el impacto económico del acceso abierto, en

particular el trabajo de John Houghton y colaboradores. El trabajo de Houghton sobre acceso abierto está orientado hacia la aplicación de técnicas de un modelo económico de forma que la inversión realizada en I+D revierta en la economía del país del que se trate, en este caso Australia (Houghton & Sheehan, 2006, 2009). Este trabajo llegó a tener una particular importancia en 2009 con la publicación ese mismo año de un informe encargado por JISC sobre los costes y beneficios del acceso abierto en la economía del Reino Unido (Houghton et al., 2009; Houghton & Oppenheim, 2010; Houghton, 2010). El informe, escrito conjuntamente con un grupo de trabajo liderado por Houghton en la Universidad Victoria de Australia y otro liderado por Charles Oppenheim en la Universidad Loughborough del Reino Unido, consistía en un detallado modelo de costes y beneficios tanto de la vía verde como de la vía dorada del acceso abierto. Se presentaban los resultados de los principales beneficios esperados de las dos vías del acceso abierto, incluso durante un período de transición, dando como resultado una mayor eficacia (costes más bajos) y al mismo tiempo aumentando sustancialmente los rendimientos de los gastos de I+D. El informe Houghton trajo consigo gran cantidad de debates y discusión en las listas de distribución, que quedó reflejado en literatura *peer-reviewed*. Posteriormente, el trabajo realizado con Alma Swan modelaba costes y beneficios durante cualquier proceso de transición al acceso abierto, argumentando que tanto la vía verde como la dorada producían beneficios económicos, aunque los mayores beneficios, al menos durante el proceso de transición, se obtendrían poniendo énfasis en la ruta verde (Houghton & Swan, 2013).

Se observa que estas áreas de interés (impacto de citación y beneficios económicos) son ejemplos de una creciente base de evidencias sobre el acceso abierto que permiten una mayor comprensión de este movimiento, aunque todavía tiene que desarrollarse y evolucionar más.

El impacto es un importante tópico unido al discurso sobre acceso abierto. Tradicionalmente fue concebido para las tasas de citación de revistas. Sin embargo, las nuevas formas de definir y medir el impacto están adquiriendo importancia en particular las llamadas métricas alternativas o

“altmetrics”, que se centran en una serie de medidas a nivel de artículo (Sud & Thelwall, 2014; Thelwall, Haustein, Larivière, & Sugimoto, 2013; Zahedi, Costas, & Wouters, 2014). Curiosamente las grandes revistas como PLOS ONE han hecho de las métricas alternativas una parte integral de sus evaluaciones de calidad (PLOS ONE, n.d.). En esta investigación doctoral se pretende ir más allá y medir el impacto de otro tipo de literatura no publicada por los canales tradicionales de publicación pero que se difunde en repositorios de acceso abierto, como es el caso de las tesis doctorales (Ferrerías-Fernández, García-Peñalvo, et al., 2015).

Las métricas tales como el uso pueden interesar fuera de la comunidad científica, donde el público puede leer y utilizar un artículo, pero no citarlo. El efecto del acceso abierto más allá de la academia, se ha establecido como el eje fundamental de apoyo para muchos mandatos de acceso abierto como se expresa en *RCUK Policy on Open Access and Supporting Guidance* (RCUK, 2013).

Según Neylon (2012) el uso de métricas como parte integral de la infraestructura de comunicación científica (en este caso evaluación de la calidad) es un ejemplo prominente de un tema de creciente importancia: la posibilidad de una literatura OA disponible en red (Neylon, 2012). Las características de esta infraestructura pueden incluir texto y datos interoperables (Shotton, 2012), que pueden ser buscados, reutilizados y recombinados. Lo que queda claro sobre una infraestructura como esta es que abarca algo más que artículos, como datos, herramientas de análisis y escenarios para la interacción que posiblemente sea lo más importante. Una cuestión clave que todo esto lleva a plantearse es cuál será la forma futura que tendrá el artículo de revista, actualmente considerado como el vehículo principal de comunicación científica. En qué grado el documento conservará su papel central y cómo se irá transformando son preguntas interesantes, cuyas respuestas surgirán con mayor claridad durante la próxima década como prevé Pinfield (2015).

I.2.4. Contenidos del acceso abierto

Cualquier tipo de contenido puede ser, en principio, de acceso abierto y puede ser digitalizado y puesto en línea sin barreras económicas o permisos para su reutilización. En ese sentido, el alcance potencial del acceso abierto es universal. Por tanto, es correcto afirmar que en algunas categorías es más fácil de aplicar y en otras más complicado en lugar de afirmar que el acceso abierto se aplica a algunas categorías o géneros y no a otros. El acceso abierto no se limita a las ciencias, donde se conoce mejor y avanza más rápido, sino que también se extiende a las artes y las humanidades. No se limita a la investigación de los países desarrollados, en los que la producción científica es más voluminosa, sino que incluye la investigación de los países en desarrollo (tampoco, por el contrario, se limita a la investigación de los países en desarrollo, donde la necesidad es más acuciante). No se limita a la investigación financiada con fondos públicos, donde el argumento es aceptado casi universalmente, también incluye la financiación privada y la investigación sin financiación. No se limita a las publicaciones presentes y futuras, en las que se centran la mayoría de las políticas de acceso abierto, sino que incluye también publicaciones anteriores. No se limita a los trabajos nacidos en forma digital, donde las barreras técnicas son menores, sino que incluye los trabajos digitalizados de material impreso, microfichas, películas y en otros soportes. No se limita a fuentes textuales, también incluye ficheros de datos, audio, vídeo, multimedia y código de programas ejecutables. (Suber, 2015)

Se han llevado a cabo campañas y prácticas exitosas para facilitar el acceso abierto a muchos tipos de contenido útil para los investigadores, entre ellos:

- Artículos de investigación revisados por pares.
- *Pre-prints* de artículos que todavía no han sido evaluadas por pares
- Tesis doctorales.
- Datos de la investigación.

- Datos gubernamentales.
- Código fuente.
- Presentaciones de congresos y conferencia (textos, diapositivas, audio, vídeo).
- Monografías académicas.
- Libros de texto.
- Novelas, cuentos, obras de teatro y poesía.
- Periódicos.
- Registros de archivos y manuscritos.
- Imágenes (obras de arte, fotografías, diagramas, mapas).
- Materiales de enseñanza y aprendizaje (“recursos educativos abiertos” y “cursos abiertos”).
- Trabajos de impresión digitalizados (algunos de dominio público, otros aún sujetos a los derechos *copyright*).

Para algunas de estas categorías, como los datos y el código fuente, se necesita que estén en acceso abierto para facilitar su comprobación y para la replicación de experimentos científicos. Para otros, como los datos, imágenes y las obras digitalizadas de otras fuentes, el acceso abierto es necesario con el fin de dar a los lectores la misma oportunidad de analizar los datos primarios de que disponen los autores. Para otras tipologías, tales como los artículos, monografías, tesis y presentaciones en conferencias y congresos, el acceso abierto permite simplemente compartir resultados y su análisis con todos los que puedan beneficiarse de ellos.

I.2.5. Políticas de acceso abierto

Los autores controlan el volumen y el crecimiento del acceso abierto. Ellos deciden si envían sus trabajos a revistas de acceso abierto (vía dorada), o si depositan sus trabajos en repositorios de acceso abierto (vía verde), y cómo hacer uso de sus derechos de explotación. Pero según Peter

Suber (2015) los investigadores todavía no están familiarizados con las opciones del acceso abierto. No tiene sentido apelar a ellos en bloque, ya que no actúan en bloque. No es difícil persuadirlos o atraerlos una vez que se capta su atención, pero son tan anárquicos y están tan saturados de trabajo y de preocupaciones que es difícil captar su atención sobre este tema. Afortunadamente, los organismos de financiación y las universidades están descubriendo sus propios intereses en el fomento del acceso abierto. Estas instituciones sin ánimo de lucro tienen entre sus objetivos promover la investigación y hacer que esta sea lo más útil y más amplia posible. Los investigadores son libres de hacer su trabajo, sin la necesidad de vincular sus ingresos a la popularidad de sus ideas. Por encima de todo, estas instituciones están en una posición única para influir en las decisiones que tomen los autores. Hoy, más de medio centenar de organismos de financiación y más de cien universidades han adoptado políticas fuertes y sólidas de acceso abierto. Cada una depende de la prioridad de las decisiones de los autores¹⁰ (Suber, 2015).

El acceso abierto es un medio para amortizar la inversión en investigación, poniendo a disposición de la comunidad científica la documentación derivada de la misma. Las políticas a favor de acceso abierto que se están desarrollando por parte de los gobiernos, instituciones y organismos de financiación nacionales y europeos contribuyen a que se multiplique el impacto de la investigación producida en los diferentes organismos científicos (Ferrerías-Fernández, 2011a).

Lo que hoy día conocemos por acceso abierto, despegó en 2001 con una carta que algunos prestigiosos investigadores dirigieron a sus colegas pidiéndoles que se negaran a publicar en revistas en las que no les concedieran el derecho a usar libremente sus artículos, al menos, en el

¹⁰La mejor lista de políticas OA de agencias financiadoras y de universidades es el Registry of Open Access Repository Material Archiving Policies (ROARMAP). <http://roarmap.eprints.org>
http://mitpress.mit.edu/sites/default/files/titles/content/openaccess/Suber_19_notes.html#note-4_2

plazo de 6 meses después de su publicación. El movimiento se consolidó y se conoce actualmente como *Public Library of Science*, <http://www.plos.org/about/what-is-plos/earlyhistory/>.

Existen políticas y declaraciones de apoyo a los procesos de implementación de los beneficios del acceso abierto para que las agencias de financiación, universidades, gobiernos y bibliotecas se sumen a las mismas. Las más destacadas son las siguientes:

- Budapest Open Access Initiative (2002)¹¹.
- Bethesda Statement on Open Access Publishing (2003)¹².
- Berlin Declaration (2003)¹³.
- Declaración de REBIUN en apoyo del acceso abierto (2004)¹⁴.
- Declaración de Barcelona (European University Association, 2008)¹⁵.
- Declaración de la Alhambra sobre acceso abierto (2010)¹⁶.
- Ghent Declaration (2011)¹⁷.
- *IFLA Statement on open access* – clarifying IFLA’s position and strategy (2011)¹⁸. IFLA Statement on Open Access (2003)¹⁹.
- SPARC Europe Statement on Open Access for EC Public Hearing on Access to and Preservation of Scientific Information (2011)²⁰.

¹¹Budapest Open Access Initiative, <http://www.soros.org/openaccess/>

¹²Bethesda Statement on Open Access Publishing, <http://www.earlham.edu/%7Eepeters/fos/bethesda.htm>

¹³Berlin Declaration, <http://oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung/>

¹⁴Declaración de REBIUN en apoyo del acceso abierto, <https://oaulpgc.wordpress.com/2009/06/22/declaracion-de-rebiun-en-apoyo-del-modelo-de-acceso-abierto/>

¹⁵Declaración de Barcelona, <http://openaccess.eprints.org/index.php?/archives/385-guid.html>

¹⁶Declaración de la Alhambra,

http://recolecta.fecyt.es/sites/default/files/contenido/documentos/declaracion_de_la_alhambra.pdf

¹⁷Ghent Declaration, <https://www.openaire.eu/en/component/content/article/223-seizing-the-opportunity-for-open-access-to-european-research-ghent-declaration-published>

¹⁸IFLA Statement on open access(2011), <http://www.ifla.org/files/hq/news/documents/ifla-statement-on-open-access.pdf>

¹⁹IFLA Statement on Open Access (2003), <http://archive.ifla.org/V/cdoc/open-access04.html>

- LERU Roadmap towards OA (2011)²¹.

A través de las mismas y mediante el establecimiento de políticas y mandatos institucionales a favor del acceso abierto se contribuye a reconocer los méritos de los investigadores que más producen, favoreciendo la excelencia, la competencia y la productividad en las entidades de I+D+i.

El principal beneficio de las políticas de acceso abierto es hacer la investigación más efectiva y sus resultados más visibles al eliminar las barreras legales, comerciales y tecnológicas de acceso a la información científica. Además, el acceso abierto previene la duplicación de esfuerzos, fomenta la transferencia tecnológica y de conocimiento y promueve la innovación y la colaboración entre diferentes grupos de trabajo.

Sin embargo, como reconocen muchos autores citados más adelante en este epígrafe, gran parte de la controversia en torno al acceso abierto se ha desarrollado como respuesta al desarrollo del acceso abierto en sí mismo pero sobre todo la causa fundamental que ha provocado dicha controversia ha sido la introducción de políticas que promueven o exigen a los investigadores poner sus resultados de investigación en acceso abierto. En los últimos cinco años, ha habido un aumento de los llamados mandatos que claramente están contribuyendo al crecimiento y auge del acceso abierto (Anne Kennan, 2011; É. Archambault, Caruso, & Nicol, 2014; Gargouri, Lariviere, et al., 2012; Kennan & Wilson, 2006; Pinfield, 2015; Xia et al., 2012).

A finales de marzo de 2016, SHERPA-Juliet²² registraba 158 mandatos internacionales emanados de organismos financiadores de investigación (Figura 10).

²⁰ SPARC Europe Statement on Open Access http://sparceurope.org/wp-content/uploads/2011/07/SPARC-Europe-Statement-on-Open-Access-Luxembourg-30-May-2011_FINAL1.pdf

²¹ LERU Roadmap towards OA, http://www.leru.org/files/publications/LERU_AP8_Open_Access.pdf

²² SHERPA-Juliet, <http://www.sherpa.ac.uk/juliet/index.php?la=en&mode=simple>

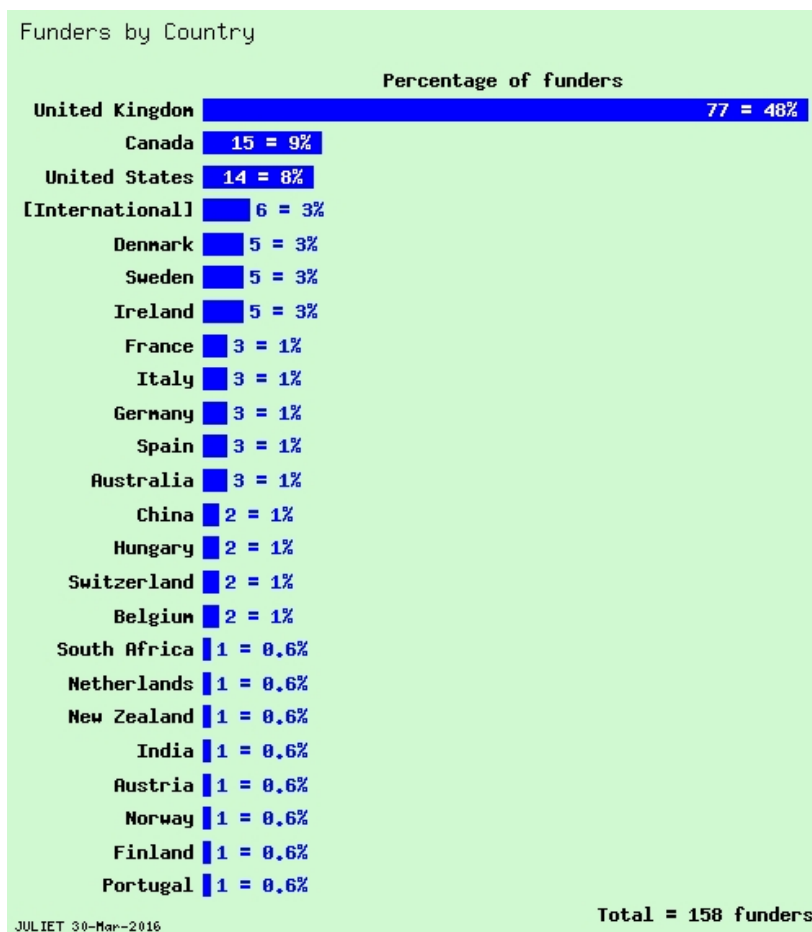


Figura 10. Organismos financiadores de investigación y mandatos (312-03-2016). Fuente: Sherpa-Juliet

ROARMap²³ identificaba 540 mandatos de organizaciones de investigación (fundamentalmente universidades). En la Figura 11 se muestra la estadística de estas políticas.

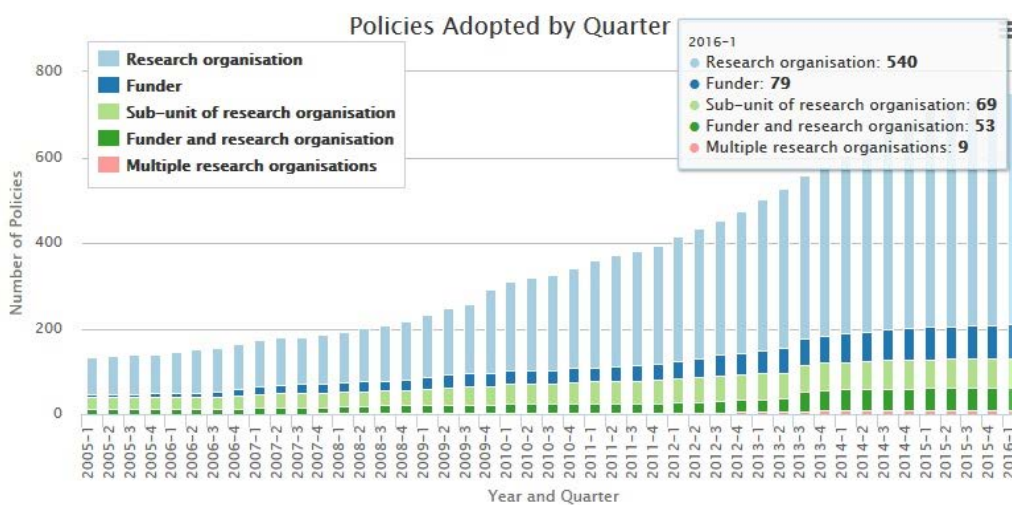


Figura 11. Gráfico de Políticas adoptadas por trimestre (31-03-2016). Fuente: ROARMAP

²³ ROARMAP, <http://roarmap.eprints.org/>

MELIBEA, <http://www.accesoabierto.net/politicas/default.php>, lista 194 mandatos de entidades financiadoras y 321 provenientes de instituciones académico/investigadoras, tal como se indica en la Figura 12.

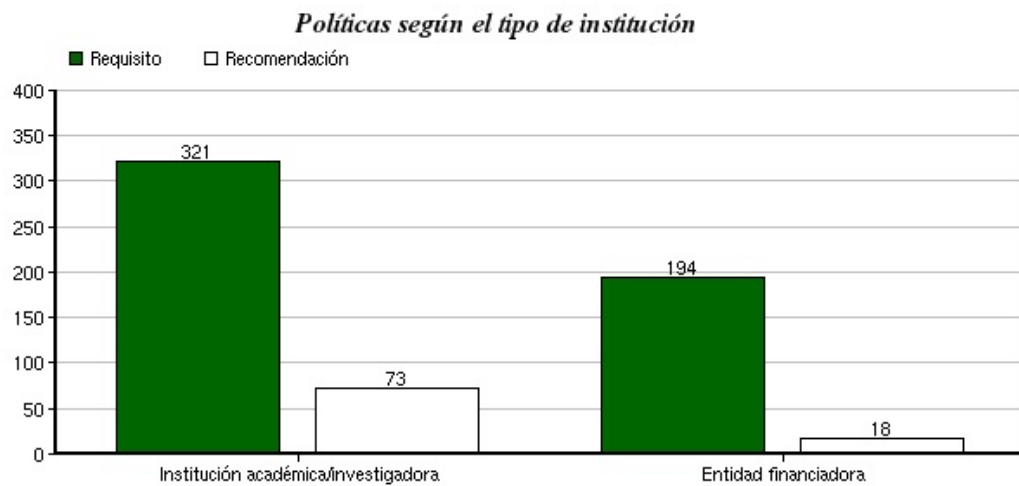


Figura 12. Políticas según el tipo de institución. Fuente: Melibea.

Vaughn (2013) proporciona un interesante relato del proceso de acuerdos llevado a cabo en EE.UU. para desarrollar y poner en marcha mandatos, proceso que ilustra la complejidad y controversia asociada con el mismo. Según Pinfield (2015) probablemente esta continua siendo la situación en 2016 donde los organismos de financiación de Estados Unidos se ven obligados a desarrollar nuevos enfoques del acceso abierto como respuesta a la *Federal Office of Science and Technology Policy*. Tales desarrollos son muy importantes, por la gran proporción de resultados de investigación mundial que estos organismos financian. Los *National Institutes of Health* de Estados Unidos fueron los pioneros en el desarrollo de políticas de acceso abierto junto con organismos como *Max Planck* en Alemania y *Wellcome Trust* en Reino Unido. En los últimos años, muchos organismos financiadores de todo el mundo han introducido nuevos mandatos o han reforzado los ya existentes con el fin de asegurar su cumplimiento como es el caso de la *Wellcome Trust* (2012). En 2012 Reino Unido puso en marcha nuevas y más robustas políticas, que incluían a los *Research Councils* (2013), el *Higher Education Funding Councils*

(2014), y las principales organizaciones médicas benéficas, a través del *Charities Open Access Fund* (Wellcome Trust, 2014).

Por otra parte, la Comisión Europea lleva impulsando políticas y medidas relacionadas con el acceso abierto desde 2006. En el 2007 se presentó a la Comisión Europea una petición para que se garantizase el acceso gratuito a los resultados de investigación financiados con fondos públicos, firmada por más de 28.000 instituciones y personas. La estrategia de la Comisión es desarrollar e implementar el acceso abierto en los resultados de investigación de los proyectos financiados por los programas marco de la Unión Europea, llamados *FP7* y *Horizon 2020*. Los requisitos del acceso abierto están basados sobre un soporte equilibrado entre el acceso abierto verde -acceso abierto inmediato o retrasado provisto por autoarchivo- y el acceso abierto dorado -acceso abierto inmediato provisto por el editor- (OpenAire, n.d.). Entre los documentos clave en el ámbito de las políticas se incluye una Comunicación de la Comisión²⁴ y las Conclusiones del Consejo²⁵ en 2007. La Comisión financia desde 2010 el proyecto OpenAIRE, (Infraestructura de Acceso Abierto para la Investigación en Europa, <https://www.openaire.eu/>) que apoya la implantación del acceso abierto en Europa suministrando una infraestructura y una red de puntos de soporte nacionales. En este momento OpenAIRE trabaja en la identificación de las miles de publicaciones asociadas a proyectos financiados por la UE, garantizando el acceso abierto a un amplio porcentaje de las mismas y la publicación en abierto de las restantes, una vez finalizado el periodo de embargo. Una segunda fase del proyecto se ocupa de asegurar el acceso a los datos primarios de investigación asociados a las publicaciones científicas. Existen ventajas significativas de tipo económico, social y educativo al hacer disponibles los resultados de investigación sin barreras financieras, legales y técnicas para su acceso.

²⁴ Comunicación de la Comisión, recolecta.fecyt.es/sites/default/files/contenido/documentos/com2007_0056es01.pdf

²⁵ Conclusiones del Consejo, recolecta.fecyt.es/sites/default/files/contenido/documentos/97236.pdf

Picarra, Angelaki, Dogan, Guy, y Artusio (2015) han escrito sobre la conveniencia de que la política de acceso abierto de *Horizon 2020* (OpenAire, n.d.) pueda ser adoptada como una política modelo en el Área de Investigación Europea (ERA), http://ec.europa.eu/research/era/optimal-circulation_en.htm, de los países para el desarrollo y aumento de alineación de las políticas de acceso. Estos autores han examinado el panorama de las políticas de acceso abierto de cinco países -Grecia, Italia, Países Bajos, Turquía y Reino Unido- con el fin de evaluar y estudiar el grado de alineación o divergencia de esas políticas con la política de *Horizon 2020* (Picarra et al., 2015). El artículo concluye considerando algunos de los impactos que puede traer la alineación de las políticas del acceso abierto y así como los mecanismos que pueden contribuir a mejorar la armonización de políticas.

En el caso de España, la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación, recoge en su artículo 37 los principales aspectos a tener en cuenta a la hora de llevar a cabo la difusión en abierto de los resultados de la investigación financiada con fondos públicos y aceptados para su publicación en publicaciones de investigación seriadas o periódicas. Las dudas surgidas en los diferentes ámbitos de aplicación motivaron la realización de un ejercicio de reflexión que permitiera ordenar cuáles son los pasos que los diferentes agentes involucrados deberían abordar para un correcto cumplimiento de la normativa legal y facilitar así la adecuada aplicación del artículo sobre "Difusión en acceso abierto", estas reflexiones quedaron plasmadas en las *Recomendaciones para la implementación del artículo 37 Difusión en Acceso Abierto de la Ley de la Ciencia, la Tecnología y la Innovación* publicadas en por Fecyt (2014).

Un estudio de Borrego (2015) mide el grado de cumplimiento con el mandato de acceso abierto del Gobierno español dos años y medio después de su implementación. Un total de 58,4% de artículos que son resultado de la financiación con fondos públicos tienen al menos una copia disponible en acceso abierto. Entre estos, el 23,8% eran acceso abierto oro, 21,8% acceso

abierto verde y 12,8% acceso abierto gris, es decir depositados en sitios web o redes sociales. La mayoría de los artículos de acceso abierto verde estaban en arXiv y en PubMed Central. Solo un 14,4% de los artículos financiados con fondos públicos estaban disponibles en repositorios institucionales, aunque más del 90% de esos artículos se debieran a proyectos llevados a cabo en instituciones que tienen un repositorio. Según el autor, existe un gran potencial para el crecimiento del acceso abierto verde, ya que más de dos tercios de los artículos que no estaban disponibles en acceso abierto fueron publicados en revistas cuyos editores permiten depositar un *pre-print* o copia *postprint*.

En cuanto al panorama español para el desarrollo de nuevas políticas de acceso abierto o simplemente la puesta en marcha de forma efectiva y eficaz de la ya existente Ley de la Ciencia, la Tecnología y la Innovación, no es muy alentador debido a la falta de auditorías o seguimiento del cumplimiento de dichas políticas por parte de los organismos financiadores con fondos públicos. Algo similar sucede, aunque en menor medida, con el mandato sobre las tesis doctorales presentadas en las universidades españolas y regulado en el Real Decreto 99/2011 por el que se regulan las enseñanzas oficiales de doctorado en España (Ministerio de Educación, 2011) y que en el artículo 14 indica que “una vez aprobada la tesis doctoral, la universidad se ocupará de su archivo en formato electrónico abierto en un repositorio institucional y remitirá, en formato electrónico, un ejemplar de la misma así como toda la información complementaria que fuera necesaria al Ministerio de Educación a los efectos oportunos”, estableciendo así un mandato de acceso abierto para las tesis doctorales. Dicho mandato se está cumpliendo en la mayor parte de las universidades de forma “flexible” como podemos observar en la normativa derivada de la Universidad de Salamanca²⁶, que se comentará más adelante.

Ha habido también controversia sobre la eficacia de los mandatos en relación con el aumento de incorporaciones al acceso abierto, pero el peso de la evidencia demuestra que las políticas

²⁶ Normativa derivada de la Universidad de Salamanca, <http://www.usal.es/webusal/node/23854>

sólidas de los organismos de financiación acompañadas por la monitorización y las sanciones dan como resultado mayores niveles de incorporaciones (Gargouri et al., 2010; Pinfield et al., 2014). También hay evidencia de la eficacia de los mandatos institucionales que también han crecido en número en los últimos cinco años (Gargouri et al., 2010; Harnad, 2009, 2011; A. Swan, Gargouri, Hunt, & Harnad, 2015). La polémica sobre el acceso abierto en estos años se ha centrado generalmente sobre las diferentes iniciativas sobre políticas y sobre los informes relacionados a menudo redactados para informar del desarrollo de la política. Quizá el más polémico de ellos fue el Informe Finch (Finch et al., 2013), informe que fue muy importante no solo por sus recomendaciones sino también por el debate internacional que generó (Andersson & Svensson, 2013; Y. Baruch, Ghobadian, & Ozbilgin, 2013; Harnad, 2013).

Según Gherab Martín (2015) el Informe Finch, publicado en junio de 2012, recomendaba priorizar la vía dorada frente a la verde y apoyarla económicamente a través de los planes nacionales de I+D del Reino Unido, empezando por el REF 2014-20206. Para ello, proponía comenzar a aplicar las recomendaciones en abril de 2013 y destinar un presupuesto para apoyar económicamente a los investigadores que publicaran en revistas doradas. El gobierno británico dio su respaldo al informe un mes después y también promulgó el uso obligatorio de la licencia *Creative Commons Attribution* o de reconocimiento de la autoría (CC-BY). El uso de esta licencia no solo permite el acceso abierto (gratuito), sino también la libre reutilización de los artículos, citando siempre la fuente. Los RCUK (Research Councils of UK) también apoyaron las recomendaciones del Informe Finch.

Algunos autores relevantes que han defendido tradicionalmente la vía verde han criticado el Informe Finch (Harnad, 2012b) ya que consideran la vía verde como la ruta más segura y rápida hacia el acceso abierto universal a la literatura científica. Stevan Harnad, máximo exponente e iniciador de la vía verde (Harnad, 2001), señala que el apoyo a la vía dorada es una victoria del *lobby* de las editoriales comerciales (Harnad, 2012b) porque, además de los ingresos por la vía

tradicional de las suscripciones, pueden obtener unos ingresos adicionales (vía autor-paga) que no se dan con la vía verde. Harnad apuesta claramente por la vía verde y aboga por no financiar a las revistas doradas y obligar a los autores a depositar sus artículos en repositorios. Otros autores de referencia, como por ejemplo Suber (2012a) o Houghton y Swan, (2013), abogan por impulsar ambas vías en paralelo.

Otro aspecto del debate sobre políticas sobre el que ha ido aumentando el interés ha sido el de licencias y permisos. En particular el debate se centra en temas sobre el grado de apertura. La distinción de Suber (2012b) entre acceso abierto “gratis” y “libre” es utilizada a menudo para enmarcar la discusión. Acceso abierto gratis quiere decir libre para leer pero acceso abierto libre significa libre para leer y reutilizar. El concepto reutilizar fue incluido en la definición de acceso abierto de la *Budapest Open Access Initiative* (BOAI, 2002) la cual incluía una larga lista de posibilidades de reutilización (“permitir a cualquier usuario leer, descargar, copiar, distribuir, imprimir, buscar”). La distinción de Suber (2012b) entre acceso abierto gratis y acceso abierto libre no son dos categorías diferentes sino dos grados de apertura. La guía de SPARC “How open is it?” incluye cuatro niveles de apertura, cada uno de los cuales corresponde con las diferentes licencias *Creative Commons* (SPARC, PLOS, & OASPA, 2012). La mayor parte del debate sobre los grados de apertura están centrados en las licencias *Creative Commons* (Carroll, 2013; Gulley, 2013; Hrynaszkiewicz, Busch, & Cockerill, 2013) con muchos financiadores de investigación que requieren licencias, tales como *CC-BY* que permitan la reutilización, incluida la explotación comercial. Algunos investigadores han manifestado sus reservas sobre que su trabajo se utilice de esta forma, a pesar de que los financiadores normalmente desean reducir la resistencia a la transferencia de conocimientos entre los resultados académicos y el uso comercial. Harnad (2012) sostiene que centrarse en la reutilización (acceso abierto libre), haciendo hincapié en la financiación del acceso abierto dorado, ha retrasado que el acceso abierto se convierta en la corriente principal de publicación.

El autor sostiene que el acceso abierto gratis verde debe ser la prioridad actual, ya que puede alcanzarse plenamente con el mínimo esfuerzo.

Puede haber diferencias significativas a la hora de implementar el acceso abierto en los diferentes países. La literatura sobre acceso abierto incluye un gran número de estudios específicos por países, incluyendo los de (W. H. Cheng, Ren, & Rousseau, 2012; R. Zhao & Wu, 2014) e India (Mukherjee, 2014; Sahu & Goswami, 2015; Sawant, 2011) que aportan estudios de investigación sobre actitudes y visiones de conjunto sobre implementación de servicios. Estos países han hecho una gran aportación a la investigación científica y a la comunicación en general a través de la literatura sobre acceso abierto (Elsevier, 2013). Los estudios generales sobre la adopción global del acceso abierto también indican una creciente adopción del acceso abierto fuera de los países occidentales (Pinfield et al., 2014; Xia, 2012).

I.2.6. Las diez claves del acceso abierto

En este “mundo” del acceso abierto, el autor sigue siendo el propietario de su obra y de sus derechos. El acceso abierto permite el libre acceso y, por tanto, gratuito a su información sin ninguna restricción. Desde la perspectiva económica, naturalmente supone un ahorro, pero además evita desequilibrios en el acceso a la información entre países ricos y pobres.

Aunque compatible con el sistema clásico de difusión de la ciencia, lo que el acceso abierto mejora es la visibilidad e impacto de los resultados de investigación y, por ende, de la transferencia de conocimiento.

Como síntesis de todo lo expuesto anteriormente es muy gráfico el decálogo de la Biblioteca de la Universidad de Salamanca (2011) que en 10 puntos sintetiza el valor del acceso abierto:

1. El acceso abierto es la difusión de forma pública de los resultados de la Investigación, a través de los repositorios en los que se depositan los documentos para su consulta libre.

2. El acceso abierto se asienta en iniciativas internacionales, a través de las cuales se acuerdan sus directrices básicas, las especificaciones técnicas y la descripción de contenidos.
3. El acceso abierto es compatible con los tradicionales sistemas de difusión de los resultados de investigación, promocionando documentos a menudo ya publicados.
4. El acceso abierto es un medio para amortizar la inversión en investigación, porque pone a disposición de la comunidad científica la documentación derivada de la misma.
5. El acceso abierto está siendo desarrollado por los organismos públicos, tomando como base que la investigación realizada con financiación pública debe tener una difusión pública.
6. El acceso abierto contribuye al crecimiento de la visibilidad de los documentos, al difundirse de forma pública y compartirse los datos de un repositorio con buscadores y recolectores.
7. El acceso abierto facilita la creación de redes científicas y grupos de investigación, ya que posibilita el conocimiento mutuo de autores que trabajan en los mismos ámbitos.
8. El acceso abierto incrementa las posibilidades de acceder a recursos de investigación sin aumentar el presupuesto disponible en una universidad o centro investigador.
9. El acceso abierto protege los derechos de los autores sobre sus obras, que se difunden libremente bajo la condición obligatoria de reconocimiento y las limitaciones que decidan.
10. El acceso abierto está siendo apoyado por políticas y normativas legales, que encuentran en este medio de difusión el mejor sistema para la transferencia de la investigación.

Capítulo I.3. Repositorios institucionales de acceso abierto

El archivo de documentos en repositorios constituye la “ruta verde” para alcanzar el acceso abierto al conocimiento. Los repositorios no “publican” documentos, hacen “pública” documentación a menudo ya publicada en revistas científicas o por los canales editoriales habituales.

Los repositorios institucionales están considerados como una de las mayores revoluciones conceptuales y tecnológicas en el campo de la publicación científica, al facilitar la diseminación rápida y generalizada de los resultados de la investigación. Pero para que los repositorios se conviertan en esa herramienta útil y poderosa, es necesario que reciban el pleno apoyo político y financiero de los gestores de las instituciones académicas, así como una amplia aceptación por parte de los académicos, los profesores universitarios y los investigadores.

El cambio experimentado en el sistema tradicional de comunicación científica ha forzado a las universidades a adoptar nuevas vías para acceder y comunicar los resultados de la investigación desarrollada por los académicos. Según el *Ranking Web* de Repositorios del Mundo (<http://repositories.webometrics.info/>) actualmente, en julio de 2016, hay casi 2.200 repositorios institucionales; dato que demuestra el interés de la comunidad académica en este tipo de comunicación científica. Sin embargo, y a pesar de que existe un gran número de repositorios que están soportados por políticas dirigidas hacia el fortalecimiento de la cooperación en la investigación científica, algunos autores como Ruiz-Conde y Calderón-Martínez (2014) opinan que en cierta medida el futuro de los repositorios es incierto.

En este capítulo se mostrará el estado del arte de los Repositorios Institucionales (RIs), se tratará el tema de su definición, contenidos y uso, además de la interoperabilidad que permite ese uso y facilita su visibilidad así como el tema de las políticas que los regulan. Además se hará

un breve recorrido por las políticas que apoyan la creación y utilización de los repositorios, políticas que son la pieza clave para conseguir el éxito de los repositorios institucionales.

El estudio de los repositorios es desde hace unos años un tema candente (Ezema, 2011) y como señala Galina (2011), en una revisión crítica de la literatura sobre la conceptualización y la función de los repositorios, existe un gran interés por ellos en la comunidad académica.

Existe una rica literatura sobre RIs, que comenzó a aparecer a principios del año 2000 (Buehler & Boateng, 2005; Crow, 2002; Lynch, 2003) y ha continuado hasta la fecha (Bhardwaj, 2014; Bhat, 2010; Bonilla-Calero, 2014; Clobridge, 2014; Connell & Cetwinski, 2010; Fan, 2015; Galina Russell, 2011; García-Peñalvo, Merlo-Vega, et al., 2010; Hawkins, Kimball, & Ives, 2013; Koler-Povh, Mikos, & Turk, 2014; Liauw, Tjiek, & Nugraha, 2011; Llorens Largo, Bayona, Gómez, & Sanguino, 2010; Marsh, 2015; Palmer, 2014; Sahu & Goswami, 2015; Schöpfel & Prost, 2013c; Stanton & Liew, 2011). Esta literatura incluye tanto estudios sobre el estado de la cuestión, como casos de estudio de experiencias individuales.

Dentro de este campo hay varias líneas de investigación, tales como las que se centran en el análisis de los factores técnicos en torno a la implementación de los repositorios (Burns, Lana, & Budd, 2013; Ezema, 2011; Ferreras-Fernández & Merlo-Vega, 2015; García-Peñalvo, Merlo Vega, et al., 2010; Giesecke, 2011; Mulhanga, Lima, Massingue, & Ferreira, 2014; Subirats et al., 2013), sobre las actitudes de autoarchivo (Carr & Brody, 2007; Singeh, Abrizah, & Karim, 2013; Xia & Sun, 2007), sobre el libre acceso, la visibilidad y el impacto (Barrueco Cruz, 2008; Davis, 2010; Fan, 2015; Galina Russell, 2011; Gaulé & Maystre, 2011; Giglia, 2010; Giusti, 2014; Kroth, Phillips, & Hannigan, 2010; López, 2013; Melero, 2007); sobre las políticas de mandato (Abadal, Ollé Castellà, Abad-García, & Melero, 2013; Ferreras-Fernández & Merlo-Vega, 2010; Unzué & Freibrun, 2015; Vincent-Lamarre, Boivin, Gargouri, Larivière, & Harnad, 2014) y sobre la

evolución de los repositorios (Keefer, 2007; Peset & Ferrer, 2008; Simpson & Hey, 2006; Sterman, 2014).

I.3.1. Los repositorios institucionales. Definición y datos

Un tema recurrente en la literatura es cómo definir un repositorio institucional y su papel en el proceso de comunicación científica. Las definiciones que se han dado de repositorio institucional han sido numerosas.

Giesecke (2011) considera que los repositorios institucionales son archivos *online* de trabajos académicos producidos localmente con el propósito de preservar y diseminar la investigación, constituyen una actividad relativamente nueva para las instituciones de educación superior. El desarrollo de los repositorios sigue el desarrollo de Internet y de la *World Wide Web*.

Crow (2002) define los repositorios institucionales como colecciones digitales que capturan y preservan los resultados de la actividad intelectual de las comunidades universitarias. Los RIs responden a dos estrategias de cara a las instituciones académicas: por una parte, aportan un componente central en la reforma de la comunicación académica a través de la estimulación de la innovación en una estructura desagregada de publicación; y, por otra parte, sirven como indicadores tangibles de calidad de las instituciones, aumentando así su visibilidad, prestigio y valor público.

Los RIs se han establecido en las organizaciones académicas y de investigación como herramientas para mostrar y facilitar la distribución en su conjunto de la investigación académica. Los RIs recogen, almacenan, diseminan y preservan recursos digitales, y muy a menudo son interoperables al utilizar *software* compatible con *Open Archives Initiative* (OAI). En muchas partes del mundo, esta infraestructura técnica es mantenida por las bibliotecas científicas y

ofrecen a las comunidades académicas un escenario adicional para participar en el acceso abierto mediante el autoarchivo (Chan, 2004).

Según Suber (2015) los repositorios de acceso abierto son colecciones *online* o bases de datos de artículos. A diferencia de las revistas de acceso abierto, los repositorios de acceso abierto no tienen su equivalente en el paisaje tradicional de la comunicación científica. Eso los hace susceptibles de ser pasados por alto o de malinterpretar su significado.

Por defecto, los nuevos depósitos en repositorios de acceso abierto son abiertos. Pero la mayoría de los repositorios que se conocen hoy en día también tienen “depósitos oscuros”, que se pueden transformar en acceso abierto pasado un tiempo. La mayoría de los repositorios de acceso abierto se pusieron en marcha para albergar artículos de investigación revisados por pares y sus *pre-prints*. Pero a menudo incluyen otros tipos de contenidos, como tesis, datos, material docente y copias digitalizadas de obras de colecciones especiales de la biblioteca de la institución. Para los académicos, los repositorios son mejores para facilitar el acceso abierto a sus publicaciones que las páginas web personales porque los repositorios proporcionan direcciones URL persistentes, toman medidas para la preservación a largo plazo y no desaparecen cuando el autor cambia de trabajo o muere (Suber, 2015).

Otros autores definen los repositorios institucionales como un conjunto de servicios para almacenar y hacer accesibles materiales de investigación en formato digital creados por una institución y su comunidad, una colección digital del producto de la investigación llevada a cabo por esa comunidad. Los repositorios institucionales se están convirtiendo en herramientas esenciales para la comunicación académica en la era digital. Los RIs pueden formar parte de un sistema mayor, nacional, regional y global de repositorios, indizados de una manera estándar y recuperable, utilizando una interfaz de acceso. También pueden proveer las bases para nuevos modelos de edición académica y pueden vincularse a otros servicios relacionados, como

enseñanza a distancia electrónica, publicación electrónica de revistas de investigación o publicación por demanda (Lynch, 2003).

McDowell (2007) considera que un repositorio debe cumplir al menos las siguientes características:

- Es un servicio institucional abierto a toda la comunidad universitaria y a todo tipo de temáticas.
- Su objetivo debe ser reunir, preservar y dar acceso a, entre otras cosas, la producción de los investigadores y docentes en múltiple formatos. Se excluiría cualquier repositorio que ponga límites a esta producción, por ejemplo, solo objetos de aprendizaje o solo de tesis, etc.
- Debe recibir contenidos de forma activa bien a través de un formulario web o simplemente a través de correo electrónico.

En esta misma línea, García-Peñalvo, García de Figuerola y Merlo-Vega (2010a) consideran que las condiciones básicas que ha de cumplir un repositorio de acceso abierto son las siguientes:

- Documentación digital: todos los documentos están disponibles para su libre acceso en formato electrónico.
- Acceso online: los documentos están almacenados en servidores accesibles a través de Internet.
- Uso público: capacidad para leer, descargar, copiar, imprimir y distribuir un documento, con la única excepción de respetar la propiedad intelectual del autor y la citación del trabajo. Los autores utilizan licencias para preservar algunos derechos (*copyleft*).
- Archivos normalizados: deben cumplimentarse estándares para la identificación digital del documento mediante el uso de protocolos internacionales.

- Colección organizada: el repositorio no es un mero depósito de documentos por lo que deben estar organizados mediante la aplicación de alguna clasificación de contenidos. Debe albergar los documentos de forma acumulativa y perpetua.

De estas definiciones se extrae como conclusión que un repositorio institucional es un conjunto de servicios prestados por las universidades y organismos de investigación, al conjunto de la comunidad, para recopilar, administrar, difundir y preservar la producción documental digital generada en la institución, cualquiera que sea su tipología, a través de la creación de una colección digital organizada, abierta e interoperable a través del protocolo OAI-PMH (protocolo para la recolección de metadatos) para garantizar un aumento de la visibilidad e impacto de la misma.

Sin embargo, entre las muchas definiciones se prefiere la que proporciona Abadal (2012) al definirlo como un sitio web que recoge, preserva y difunde la producción académica de una institución permitiendo el acceso a los objetos digitales que contiene y a sus metadatos.

Los contenidos fundamentales de los repositorios son las publicaciones que se derivan de la investigación (artículos de revista, informes de investigación, congresos, tesis doctorales, etc.), aunque en muchos de ellos se puede encontrar documentación académica (material docente, actividad institucional, etc.) e incluso documentación y fondo patrimonial como es el caso del repositorio Gredos (material de archivo, fondo antiguo de la biblioteca digital, etc.).

Los objetivos principales de los repositorios son los de favorecer la difusión de los contenidos académicos de la institución a la que sirven o de la disciplina a la que se dedican, dar visibilidad a la investigación realizada por la institución y sus miembros y facilitar la conservación y preservación de los documentos generados por una institución. Un repositorio institucional es sobre todo, además de un depósito donde almacenar y preservar documentos digitales, la imagen de la producción científica y académica de su institución.

Existen tres formas de incorporar los contenidos a los repositorios:

1. Fundamentalmente mediante el autoarchivo, es decir, que los depósitos los realizan los propios autores y los metadatos son revisados por los gestores del repositorio y bibliotecarios.
2. También existe el depósito delegado, que se lleva a cabo por parte de los gestores del repositorio.
3. Y mediante la carga masiva, que se realiza a través de la recolección de contenidos procedentes de revistas o de otros repositorios.

En todos los casos y como buenas prácticas, habrá que relacionar los documentos depositados con la fuente original donde se han publicado.

Como afirman algunos autores, los repositorios cuyas funciones principales desde un principio eran las de estar destinados a proporcionar un depósito (facilitar el autoarchivo para preservar el legado académico) y a facilitar el acceso (facilitar los procesos de recuperación de la información), se han convertido en un componente clave del acceso abierto al conocimiento (Orduna-Malea & Delgado López-Cozar, 2015) y constituyen una importante parte de la implementación del acceso abierto desde el principio del movimiento (Björk et al., 2014; Pinfield, 2015).

Los directorios de repositorios son una excelente fuente de información para conocer los archivos abiertos de las organizaciones. Los directorios más completos son:

- Directory of Open Access Repositories (OpenDOAR), <http://www.opendoar.org/>.
- Registry of Open Access Repositories (ROAR), <http://roar.eprints.org/>.
- Hispana (directorío y recolector), <http://hispana.mcu.es/>.
- BuscaRepositorios, <http://www.accesoabierto.net/repositorios/>.

Un análisis del crecimiento global de los repositorios desde 2005 hasta 2016 utilizando los datos de OpenDOAR reporta un aumento exponencial en el número de repositorios que creció de 128 en diciembre de 2005 hasta a 3.090 en mayo de 2016. Los datos de la evolución mundial de los repositorios que ofrece el directorio OpenDOAR se resume en la Figura 13.

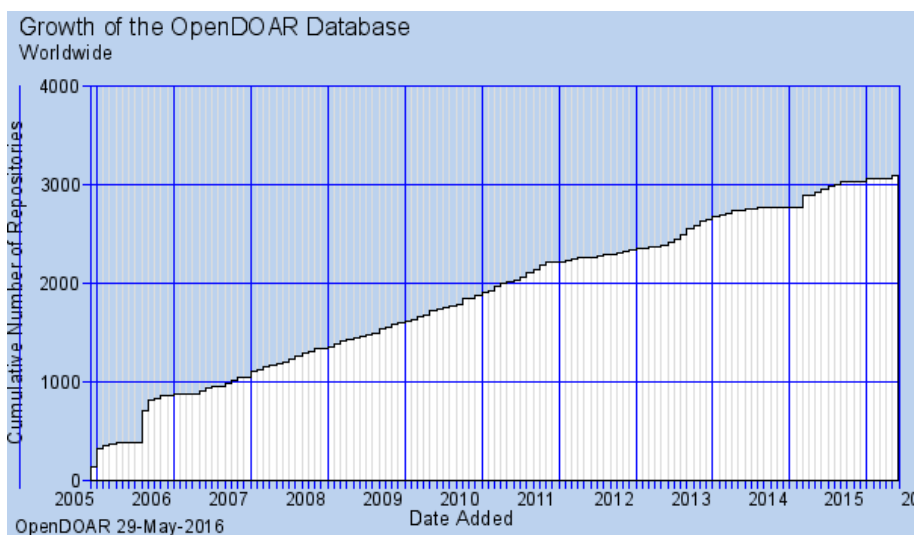


Figura 13. Crecimiento mundial de Repositorios. Fuente: OpenDOAR

Se observa que el número total de repositorios registrados en el directorio OpenDOAR asciende actualmente (mayo de 2016) a un total de 3.018 repositorios.

En cuanto a los repositorios institucionales registrados en OpenDOAR la cifra asciende a 2.612 RIs, en mayo de 2016, en la Figura 14 se muestra la evolución de los mismos.

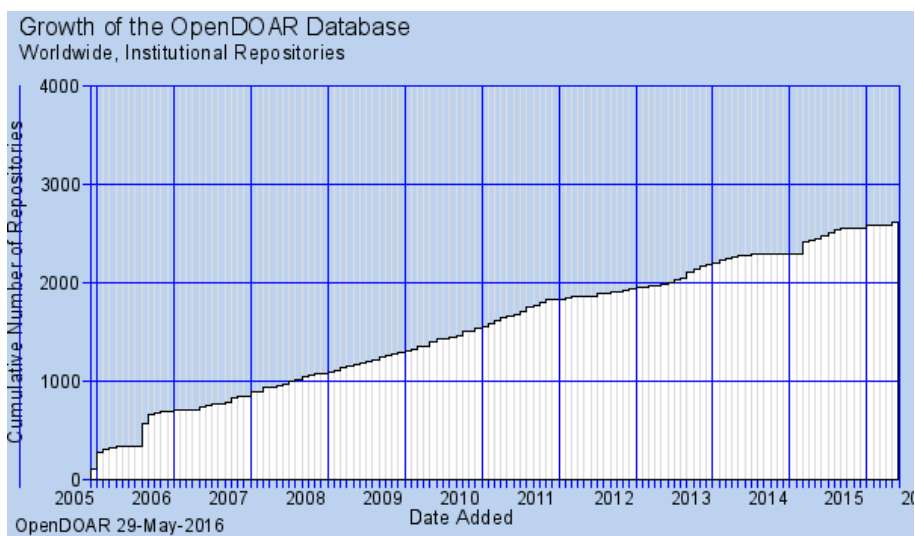


Figura 14. Crecimiento mundial de Repositorios Institucionales (RIs). Fuente: OpenDOAR

Implementar un repositorio institucional requiere un considerable esfuerzo de planificación y de compromiso. Una de las principales motivaciones para su creación es la de permitir el acceso abierto a los resultados de la actividad científica y académica de la institución. Pero quizás esta razón solamente sería insuficiente para justificar el esfuerzo considerable que supone la creación de repositorios.

En un primer nivel, un repositorio institucional es el reconocimiento de que la actividad intelectual de nuestra universidad estará representada cada vez más en soporte digital y que la principal responsabilidad de esta es ejercer el control sobre su producción intelectual, haciéndola accesible, fácilmente recuperable y asegurando su permanencia en el tiempo.

Las Instituciones de Educación Superior tienen que gestionar sus contenidos digitales de forma más efectiva y transparente que antes. En Europa se están desarrollando actividades (EEES) que ofrecerán beneficios tangibles a aquellas instituciones que puedan demostrar y explotar sistemas y estrategias de información efectivas.

La gran ventaja que proporcionan estos sistemas es que ayudan a la institución a desarrollar estrategias coherentes y coordinadas para la captura, identificación, almacenamiento, conservación y recuperación de sus contenidos digitales. El tratamiento gestionado de estos contenidos aumenta las oportunidades para un uso más efectivo de los resultados de la actividad de la institución y estimula la colaboración entre las diferentes disciplinas y comunidades. La posibilidad que ofrecen estos repositorios de reutilizar los contenidos abre un amplio abanico de aplicaciones, por ejemplo en el ámbito de los objetos de aprendizaje. Por otra parte, suponen un medio de romper el ciclo de depósitos o almacenes individuales de conocimiento dentro de la institución, ofreciendo un espacio común de almacenamiento con acceso para todos (Ferrerías-Fernández, 2010c).

Por otra parte, en cuanto a la gestión adecuada de los recursos de aprendizaje mejoraría significativamente la reutilización de los mismos, y conduciría a un considerable ahorro de tiempo y esfuerzo en la generación de materiales didácticos. Sin embargo, los recursos disponibles a través de los denominados *Learning Management Systems* (LMS), como por ejemplo *Moodle* o *WebCT*, no facilitan la interoperabilidad ni la reutilización de los recursos para el aprendizaje. Una solución para generar un ecosistema de e-learning es la creación de Repositorios de Objetos de Aprendizaje Reutilizables (ROAR) o Repositorios Institucionales (RI) que son bases de datos con servicios de captación, almacenamiento, indexación, preservación y redistribución abierta (open access) de contenidos educativos y de investigación en formato digital pertenecientes a una comunidad universitaria (Valverde Berrocoso, 2013).

Como indican Llorens Largo (2011) e Illanas y Llorens Largo (2011), los repositorios institucionales existen para almacenar y mantener información digital que a dar visibilidad a todo el contenido docente o científico que se genera dentro de la institución. Los mismos autores señalan que estas herramientas no son nada sin la colaboración del profesorado ya que son ellos los que deben dotarlas de contenido, y de esta forma, a través de iniciativas de este tipo, será posible aprovechar recursos educativos digitales, de calidad, gratuitos y en abierto. Se trata de desarrollar al máximo las potencialidades comunicativas a través de las herramientas que proporciona la red, que permiten la interacción y el intercambio ágil de información; comunicar, compartir y colaborar se convierten en la clave de estas nuevas utilidades y servicios. La combinación de todas estas herramientas está imponiendo una forma diferente de relación, tanto dentro de la propia universidad, como con el entorno.

En definitiva, un repositorio institucional ofrece la difusión más amplia posible de toda la oferta de la producción intelectual digital generada en una institución, incrementando la visibilidad y el prestigio de la misma y demostrando su valor para las fuentes de financiación y subvenciones. Además puede aumentar la visibilidad de sus investigadores, ampliando la difusión y el uso de

sus trabajos; puede estimular la innovación, facilitar un análisis cualitativo del trabajo de sus miembros, apoyar las tareas de enseñanza y aprendizaje, servir de sistema de registro de ideas y ofrecer un catálogo del capital intelectual de la institución. Los repositorios institucionales reconocen y dan cabida a los objetos digitales científicos no incluidos en los canales tradicionales de publicación. Finalmente, los repositorios institucionales pueden mejorar la comunicación científica y hacer avanzar la investigación permitiendo a los usuarios localizar y recuperar información relevante más rápida y fácilmente.

Ante esta realidad es una prioridad de las instituciones, en este caso de la Universidad, tomar conciencia de la necesidad de conservar y hacer accesible a largo plazo el contenido digital de los repositorios institucionales, como reflejo que son de la actividad investigadora y docente de la Universidad, a la vez que depositarios del fondo patrimonial de la misma.

I.3.2. Uso del repositorio y captación del contenido

Casi todos los autores que han escrito sobre el autoarchivo en los repositorios refieren problemas a la hora de fomentar la participación de los autores y el uso del repositorio. En este sentido Llorens Largo (2011) refiere el problema, por una parte, a la falta de comprensión de las ventajas de colaborar y compartir, ya que todavía se conservan ciertas reticencias “a soltar aquello que consideramos nuestro” y, por otra parte, a la alarma generada en algunos docentes e investigadores al confundir esta filosofía con una descontrolada libertad, de gratuidad y de libre disposición del trabajo ajeno y pérdida de la autoría de lo que se comparte abiertamente.

Por otra parte, la evaluación de Davis y Connolly (2007) demostró que el repositorio de la Universidad de Cornell tenía poco contenido y era poco utilizado por el personal académico. Las razones clave de la falta de uso incluyen la preferencia por otras alternativas a los repositorios, la percepción de que los repositorios eran redundantes, las dificultades técnicas, preocupación por el posible plagio de sus trabajos, preocupación referente a la calidad y estatus del repositorio y

preocupación los derechos de autor. Otra investigación de Foster y Gibbons (2005) para conocer la opinión de los profesores con el fin de mejorar el repositorio aumentando los contenidos de este, encuentra que la mayoría de los investigadores no percibe ningún beneficio potencial con el uso del repositorio. La aprehensión hacia el depósito en el repositorio y hacia la publicación de acceso abierto en general, parece centrarse en torno a tres cuestiones fundamentales: la falta de motivación para la autoarchivo; las preocupaciones en torno a la propiedad intelectual, los derechos de autor y el plagio; y muestran actitudes negativas hacia la publicación y el archivo en acceso abierto como modos legítimos de la comunicación científica.

Se han ido adoptando políticas de mandato de depósito en acceso abierto por muchas instituciones, en un principio para impulsar el contenido del repositorio y crear una colección sostenible y accesible de los resultados de investigación (Sale, 2006). Estas políticas de mandato se han venido aplicando a tipos específicos de resultados de la investigación, a resultados del personal académico, o a las tesis doctorales. Según Sale (2006) aunque los mandatos tardan tiempo en ser integrados en los procesos de trabajo de los académicos han demostrado ser una efectiva vía para lograr el crecimiento y garantizar la sostenibilidad de las colecciones depositadas. Sin embargo, algunos desarrolladores de repositorios han considerado que los mandatos dañan la imagen del repositorio como servicio integral a la comunidad académica y es más importante la visión del personal académico que la garantía de asegurar el regular crecimiento de contenido del repositorio (C. Palmer, Teffeu, & Newton, 2008).

Según Suber (2012b, 2015) la razón por la que la tasa de depósito voluntaria es inferior a la tasa por mandato no suele ser la reticencia al acceso abierto en sí. Casi siempre es la falta de familiaridad con el acceso abierto verde, debida a la creencia de que todo el acceso abierto es dorado, a los malentendidos en torno al acceso abierto verde, debidos a la creencia de que viola los derechos de autor, de que no pasa por la revisión por pares, o de que excluye la posibilidad

de publicar en una revista respetable, además del temor a que sea una pérdida de tiempo. En este sentido, el desconocimiento y los malentendidos por parte de los autores son los mayores obstáculos para el acceso abierto que lo puedan ser la oposición de ellos mismos o de las editoriales.

Como se observa en la Figura 15 la mayoría de los repositorios son institucionales (84,5 %), con una proporción mucho más pequeña de otros tipos de repositorios particularmente los repositorios temáticos (9,6). Sin embargo los repositorios temáticos son los responsables de la aportación mayor de contenido a los repositorios de acceso abierto. Algunos de los repositorios temáticos más grandes como arXiv y PubMed Central, son ahora servicios consolidados (Nicholas et al., 2012). Sin embargo, hoy en día muchos de los RIs existentes son todavía pequeñas implementaciones piloto.

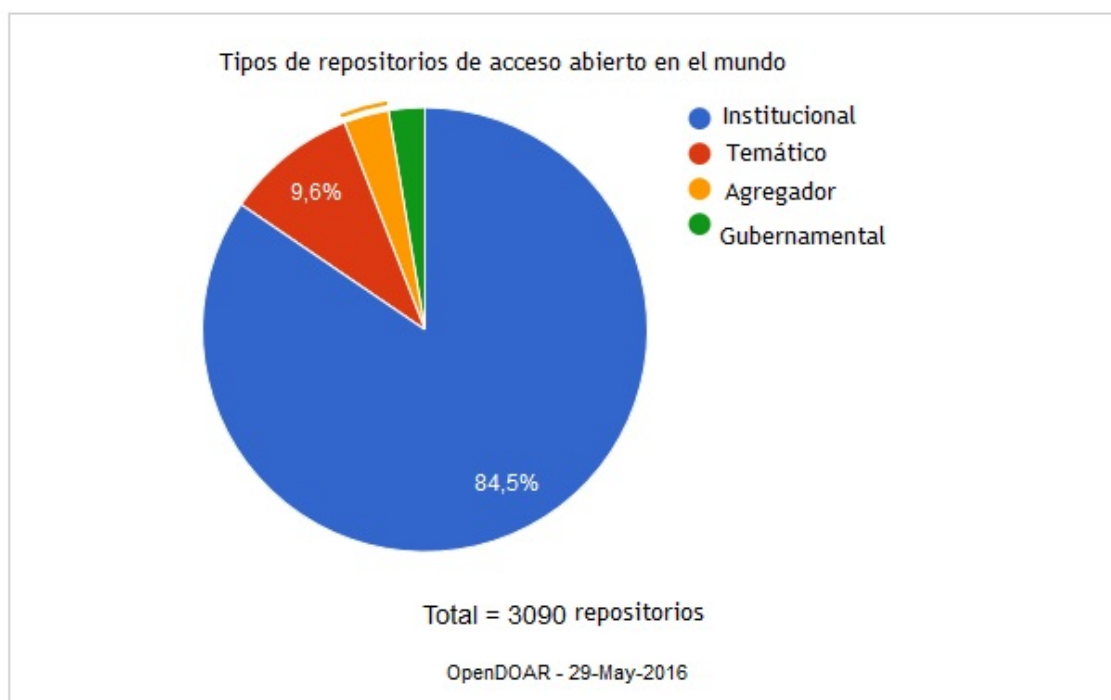


Figura 15. Tipos de repositorios de acceso abierto en el mundo. Fuente: OpenDOAR

En relación a la aportación de contenidos en repositorios ha surgido un problema en los últimos diez años que es el depósito con embargo (Laakso, 2014; Sutton, 2013). Sutton (2013) ha estudiado este cambio hacia los embargos observando que son más largos y más restrictivos

que antes, con reglas complejas sobre dónde deben depositarse los ítems (por ejemplo, en páginas web personales pero no en repositorios) y por qué (por ejemplo voluntariamente pero no en respuesta a un mandato) así como cuándo (por ejemplo después de 12 meses o más).

Mientras que alimentar y mantener la red mundial de repositorios OA puede seguir siendo un desafío, las discusiones técnicas han evolucionado a partir de las cuestiones fundamentales de la creación de repositorios conectados a través de protocolos de interoperabilidad (Lagoze & Van de Sompel, 2003) para la integración de repositorios en una infraestructura académica más amplia que pueda mejorar la investigación y la gestión de la investigación. Un ejemplo de esto es la implementación de estándares como ORCID, identificador de autor. También existe un foco de estudio sobre la usabilidad diseñada para ayudar a reactivar los repositorios con el fin de trabajar a escala (R. Johnson, 2015).

I.3.3. Factores de éxito de los Repositorios Institucionales

A pesar de que las universidades y otras organizaciones de investigación en todo el mundo gastan grandes cantidades de dinero para crear bibliotecas digitales y repositorios institucionales (Tripathi & Jeevan, 2011), los repositorios han tenido a lo largo de su corta historia bajas tasas de éxito.

McGovern (2009) relaciona tres posibles indicadores para el éxito de un RI: envíos de contenido, uso, y soporte, siendo el contenido el requisito previo para el establecimiento de un RI bien poblado y ampliamente utilizado. McGovern insiste en la necesidad de una definición explícita para el éxito de un RI y sugiere que podrían establecerse incentivos apropiados para conseguir lo que debería ser el mandato esencial de los RIs, esto es capturar y preservar el contenido con el fin de optimizar el valor e impacto de los activos digitales a través del tiempo. Con el crecimiento actual e interés en esta infraestructura, es importante tener en cuenta los factores que afectan el desarrollo y mejora de los RIs.

Westell (2006) sugirió ocho factores de éxito para la evaluación de los RIs en Canadá, comprendiendo seis factores internos (mandato, integración con planificación, modelo de financiación, medida, promoción y estrategia de preservación) y dos factores externos (relación con los centros de digitalización e interoperabilidad).

Thibodeau (2007) propuso un marco mucho más general para la evaluación de los RIs, articulado a lo largo de cinco puntos: servicio, orientación, cobertura, colaboración y estado.

Swan (2007) recomendaba un marco de trabajo de calidad basado en cuatro aspectos: captación de contenido, sensibilización y participación de los usuarios, prácticas de flujo de trabajo y disciplina financiera. Por su parte, Cassella (2010) recomendaba un conjunto de catorce indicadores para medir el coste-eficacia de los repositorios y el éxito basado en cuatro perspectivas: usuario, proceso interno, financiación, y aprendizaje y crecimiento.

Los factores de éxito relativos a la tecnología en los RIs están asociados con aspectos como el *software*, la usabilidad, las colecciones digitales, el rendimiento, la infraestructura técnica, y la interoperabilidad. La usabilidad del *software* tiene impacto sobre la experiencia de usuario en repositorios digitales y afecta al éxito del RI (Z. Deng & Li, 2008) y a los diferentes marcos de trabajo de evaluación de la usabilidad que han sido desarrollados para los RIs (Kim & Kim, 2008). El funcionamiento de los repositorios digitales es crucial desde tres perspectivas: la tecnología del sistema, los métodos de recuperación de la información, y los servicios personalizados para usuarios (Zhao, Niu, Cao, & Dai, 2010).

La dimensión "persona" en los RIs implica cultura organizacional, apoyo de la dirección, política, satisfacción del usuario, y personal. La cultura organizacional afecta el crecimiento y el uso de los RIs (Shearer, 2003), juega un papel importante en el desarrollo de percepciones de grupo, en la coordinación de las actividades de grupo y en la toma de decisiones, y en el balance de intereses individuales y de la institución. La institución se encuentra en la cúspide del soporte de

gestión en áreas como política, financiación, y preservación digital. El apoyo a la gestión en áreas como las políticas, financiación y la preservación digital es importante para mostrar el nivel de compromiso con la implementación del RI. La institución deberá tener una política y un detallado plan que recoja todos los aspectos necesarios para conseguir el éxito de implementación de los RIs, incluido el personal, la comunicación y los planes de formación y los resultados para la planificación. Formar personal experto y asegurar que el personal esté familiarizado con la teoría y la práctica de los RIs son elementos vitales para el éxito (Cullen & Chawner, 2010).

El éxito relacionado con los servicios del RI implica un servicio de descripción (Lynch, 2003) que incluya soporte técnico, preservación y servicio de metadatos, derechos de autor y licencias, compartir recursos, y mantenimiento.

Algunos autores sostienen que el éxito de los RIs se encuentra en los servicios de valor añadido que apoyan activamente el proceso de la comunicación académica, incluyendo el registro de la idea intelectual, que certifica la calidad y validez de los resultados, difundir la investigación a los usuarios, y preservar el registro académico para el futuro (M. Ramírez & Parham, 2010). La prestación de una gama completa de servicios de apoyo académico y de investigación es una de las recomendaciones para la implementación exitosa de un RI (Jain, 2011). Los servicios son un factor interno del éxito de los RIs (Markey, Rieh, St. Jean, Yakel, & Yao, 2009). Tripathi y Jeevan (2011) sugieren que los usuarios deben ser formados en cuestiones de propiedad intelectual de esta manera los servicios de RI podrán alcanzar el éxito gracias a ser compatibles con los derechos de autor. Para Thibodeau (2007) el éxito de los repositorios vendrá determinado en último término por el uso de la colección, igualmente Sawant (2011) indica que los usuarios son uno de los factores más importantes de la supervivencia a largo plazo de los RIs. El nivel de satisfacción de los usuarios tendrá un impacto directo en el éxito de los RIs (Z. Deng & Li, 2008). Dörner y Revell (2012) señalan que para lograr el éxito los RIs necesitan a sus clientes tanto

para depositar documentos como para acceder y usar esos documentos. Galina Russell (2009) demostró en su tesis doctoral que los gestores de los RIs consideran importantes los datos de uso, especialmente para promocionar el repositorio y para asegurar la financiación del mismo.

La colección es el núcleo de los servicios de un RI. Los autores citan a menudo la captación de contenido como uno de los factores de éxito en el desarrollo de un RI (Bell, Foster, & Gibbons, 2005; Ferreira, Rodrigues, Baptista, & Saraiva, 2008; Shearer, 2003). La investigación preliminar de Cullen y Chawner (2010) sobre la construcción de repositorios institucionales desde la perspectiva de las bibliotecas consideran que el crecimiento y el uso del repositorio junto al tamaño de la colección son indicadores de éxito.

Métricas tales como el número de ítems descargados indican que el contenido del repositorio ha sido localizado y usado. Deng y Li (2008) sugieren que el control de calidad de una colección digital debería abarcar el alcance, la autoridad, precisión, y el *copyright*. Además las colecciones digitales deberían ser capaces de satisfacer las necesidades de los usuarios.

El *Center for Research Libraries* (2007) remarcó que los pequeños éxitos de los RIs en el mundo eran dignos, sin embargo no dan lugar a medidas que podrían utilizarse para evaluar beneficios para las instituciones relacionados con la inversión de tiempo y energía ya sea en relacionados con la captación y contribución a la comunicación científica o con criterios más técnicos adoptados como estándares. De hecho, no hay acuerdo en la literatura acerca de cuál de los muchos factores de éxito citados son cruciales para todos los RI. Ninguna investigación previa ha desarrollado un conjunto de indicadores aceptados de éxito para los RI (Lagzian, Abrizah, & Wee, 2015).

Lagzian, Abrizah y Wee (2015) realizaron un análisis de brechas para medir la importancia percibida y el rendimiento real de los repositorios institucionales. Los autores analizaron los factores críticos de éxito en la implantación de los RIs. Tuvieron en cuenta las perspectivas de

los gestores de repositorios sobre la manera en que perciben la importancia de los factores identificados como críticos y considera el rendimiento real de los RI en relación con esos factores. El análisis de brechas compara dos puntos de vista sobre la implementación de los RIs dónde están los repositorios institucionales (situación actual) y dónde quieren estar (importancia percibida). Los seis factores críticos de éxito identificados por este estudio fueron la gestión, los servicios, la tecnología, la práctica de autoarchivo, las personas y los recursos.

Los repositorios institucionales se han convertido en un vector significativo de la comunicación científica. Actualmente, cuatro repositorios sobre cinco son repositorios institucionales. Una de sus características reside en su gran diversidad (Schöpfel & Prost, 2013).

Los repositorios institucionales de acceso abierto han cumplido ya 15 años de existencia, al igual que la definición correspondiente para OAI-PMH, que fue lanzada en 2001. Hasta la fecha, hay más de 3.000 repositorios en todo el mundo, con una cobertura colectiva de más de 100 millones de objetos de diferentes tipos de contenido y calidad. Durante los primeros años, el principal foco de atención de los repositorios institucionales era proporcionar una ventana local al contenido producido en la institución, con servicios desarrollados en torno al apoyo de esta perspectiva institucional. Los esfuerzos de interoperabilidad estaban centrados en el intercambio de metadatos vía OAI-PMH, y en servicios de terceros que fueron establecidos principalmente para agregar metadatos de repositorios o exponer las páginas web de los repositorios para poder ser indexados por los motores de búsqueda²⁷.

²⁷ OpenDOAR. www.opendoar.org 3238, BASE www.base-search.net 4.693, en septiembre 201

I.3.4. La interoperabilidad y visibilidad de los repositorios

I.3.4.1. La interoperabilidad y los repositorios. Definición y datos

Los sistemas de información Web se encuentran en un punto en el que ciertas tecnologías y prácticas han alcanzado un gran nivel de madurez en su desarrollo y han sido adoptados por parte de la industria. Estos sistemas de información se han extendido a las instituciones en diferentes contextos y ámbitos, pero en un mundo global, estos sistemas necesitan ser interoperables a través de estándares que definen formas de distribuir contenidos o servicios de alta calidad (Alier Forment et al., 2012, 2010; García-Peñalvo, Alier Forment, & Lytras, 2012).

Se entiende por interoperabilidad la capacidad de los sistemas, servicios y organizaciones para trabajar juntos de un modo transparente hacia objetivos comunes o diferentes. La interoperabilidad tiene que ver principalmente con el contexto de la búsqueda y acceso a los recursos. La interoperabilidad es un término usado para describir diferentes modelos de intercambio de datos entre sistemas de información, basada en estándares y protocolos.

La interoperabilidad permite el intercambio efectivo de información entre cuerpos interoperables permitiendo que se generen nuevos conocimientos (Miller, 2000). Los repositorios pueden ser accesibles abiertamente e interoperables con otros repositorios, preferiblemente usando el protocolo *Open Archive Initiative Protocol for Metadata Harvesting* (Chan, 2004). Es crucial, por tanto, que la interoperabilidad continúe siendo una preocupación principal, en el ámbito internacional, que permita la recolección y posterior reutilización del contenido (McCulloch, 2006). La interoperabilidad garantiza que cualquier usuario en cualquier parte del mundo pueda buscar archivos en los repositorios también ubicados en cualquier parte (Hanief Bhat, 2010).

Según Miller (2000) la interoperabilidad puede alcanzarse por cuatro vías: a través de productos de ingeniería, la industria, el acceso a la tecnología, y la implementación de estándares. El

mismo autor afirma que para ser interoperable, hay que comprometerse de forma activa en un proceso constante para asegurar que los sistemas, procesos y cultura de la organización sean gestionados de tal forma que se maximicen las oportunidades para intercambiar y reutilizar la información, tanto interna como externamente.

Reitz (n.d.) define la interoperabilidad como la capacidad de los sistemas *hardware o software* para comunicar y trabajar eficazmente con otros sistemas en el intercambio de datos, usualmente un sistema de diferente tipo, diseñado y producido por diferente vendedor.

En *New IEEE Standard Dictionary of Electrical, Electronics and Computer Terms* (1994) se define la interoperabilidad como la capacidad de dos o más sistemas o componentes para intercambiar información y para usar la información que ha sido intercambiada.

Según Johnson (2002) la interoperabilidad es la capacidad del sistema de repositorios para proporcionar acceso a su contenido a través de múltiples motores de búsqueda y otras herramientas de descubrimiento de datos. De esta forma, cuanto más visible es el contenido del repositorio más interoperable es. Warner (2003) recomienda *Open Archive Initiative* (OAI) como infraestructura para repositorios, ya que evita que los repositorios individuales se conviertan en islas aisladas de información.

Seadle (2008) también recomienda planificar la interoperabilidad para que los documentos puedan pasar fácilmente de un sistema a otro. Nicholson y Shiri (2003) están a favor de la planificación de la interoperabilidad más allá de las agrupaciones de bibliotecas digitales existentes, ya que el OAI-PMH proporciona una forma estandarizada y eficiente para que se comuniquen dos computadoras o sistemas en red con el fin de compartir metadatos. Uno de esos sistemas se designa proveedor de datos o repositorio y el otro como proveedor de servicios o recolector o *harvester*. El OAI-PMH no solo mejora la interoperabilidad, sino que también permite a los repositorios compartir recursos y mejorar así el acceso a los usuarios.

Para Hakimjavadi y Masrek (2013), en el campo de los repositorios institucionales, el término interoperabilidad representa a menudo un método específico de interoperabilidad como por ejemplo la recolección de metadatos mediante OAI-PMH.

Así pues, una característica primordial de los repositorios es que sean interoperables y el grado de visibilidad del contenido de los repositorios está estrechamente relacionado con la interoperabilidad de los mismos (Hanief Bhat, 2010). Los diferentes repositorios institucionales siguen protocolos o estándares técnicos que posibilitan su consulta desde el mismo portal del repositorio de forma individual, o a través de otros portales o sitios como son los recolectores que actúan como metabuscadores de documentación científica. Por ello, el depósito en repositorios de acceso abierto contribuye a un mayor conocimiento de las publicaciones, a un aumento de los índices de citaciones y a la mejora de los resultados de transferencia de resultados de la investigación de los organismos científicos (Ferrerías-Fernández & Merlo-Vega, 2015).

La normalización en cuanto a la descripción de los documentos digitales y la interconexión de servidores permite que los datos de los documentos depositados en los repositorios científicos se localicen desde buscadores y plataformas de recuperación de información (García-Peñalvo, Merlo-Vega, et al., 2010; Ríos-Hilario, Ferreras-Fernández, & Martín-Campo, 2013; Ríos-Hilario et al., 2012). Los documentos científicos se depositan en un repositorio concreto pero los datos descriptivos y de localización se comparten, de forma que la visibilidad de un documento en acceso abierto se multiplica exponencialmente, así como el posible impacto de sus contenidos (Ferrerías-Fernández, 2015; Ferreras-Fernández, Merlo-Vega, et al., 2015; Setenareski, Sunye, & Shima, 2013).

I.3.4.2. La Iniciativa de Archivos Abiertos. El protocolo OAI-PMH

La interoperabilidad se apoya, entre otros, en los estándares abiertos para la comunicación entre sistemas y para la descripción de recursos y colecciones. La Iniciativa de Archivos Abiertos (OAI) pretende desarrollar y promocionar normas de interoperabilidad que tienen como objetivo el facilitar la difusión eficaz de los contenidos. El objetivo esencial de los archivos abiertos es permitir el acceso a los materiales en la web por medio de repositorios que, interoperando unos con otros, permitan el intercambio de metadatos, su almacenamiento y publicación. La necesidad de la comunidad de *e-print* (ediciones preliminares electrónicas) de obtener una solución de interoperabilidad de bajo nivel para el acceso entre repositorios heterogéneos dio lugar a la creación de la *Open Archives Initiative* (OAI). En el apartado misión de la OAI se dice: “*Open Archives Initiative* desarrolla y promueve normas de interoperabilidad que tienen por objeto facilitar la difusión eficaz de los contenidos”²⁸.

De esta estrategia de archivos abiertos se podrán beneficiar muchas comunidades. Internet y la gran cantidad de documentos en formato digital han aumentado el número de usuarios potenciales de los repositorios. Los documentos pueden ser más accesibles y además utilizados con distintos objetivos de los que motivaron inicialmente la creación de los repositorios. Por otra parte, la posibilidad de acceder a múltiples repositorios permite la construcción de nuevos tipos de servicios que puedan atender mejor las necesidades de los usuarios.

En relación con la OAI, la recolección se refiere específicamente a la recogida de los metadatos de una serie de repositorios distribuidos en un almacén de datos combinados. Y la aplicación cliente que suministra las peticiones OAI-PMH sería lo que se denomina un recolector o *harvester*.

²⁸ What is the mission of the Open Archives Initiative?, <http://www.openarchives.org/documents/FAQ.html#What%20is%20the%20mission%20of%20the%20Open%20Archives%20Initiative>

El Protocolo OAI para la Recolección de Metadatos (OAI-PMH), define un mecanismo para la recolección de registros que contienen los metadatos de los repositorios. Ofrece a los proveedores de datos una opción técnica sencilla para poner sus metadatos a disposición de servicios basados en los estándares abiertos HTTP (*Hypertext Transport Protocol*) y XML (*Extensible Markup Language*). Los metadatos que son recolectados pueden estar en cualquier formato establecido por una comunidad (o por cualquier conjunto específico de proveedores de datos y proveedores de servicios), con independencia de que hayan establecido el Dublin Core no cualificado para proporcionar un nivel básico de interoperabilidad. Así pues, los metadatos de distintas fuentes pueden ser reunidos en una base de datos, y se pueden ofrecer servicios sobre la base de esta recolección centralizada, o agregación de datos. El vínculo entre estos metadatos y el contenido no está definido por el protocolo OAI. El protocolo OAI-PMH no proporciona una búsqueda a través de estos datos, simplemente permite reunir los datos en un sitio. Para ofrecer un servicio, el método de la recolección debe combinarse con otros mecanismos.

Es conveniente aclarar los conceptos proveedores de datos y proveedores de servicios. Los proveedores de datos manejan el depósito y la publicación de los recursos en un repositorio y exponen los metadatos de los recursos del repositorio para que puedan ser recolectados. Ellos son los creadores y conservadores de los metadatos y de los repositorios de recursos. Los proveedores de servicios recolectan los metadatos de los proveedores de datos y, emplean los metadatos recolectados con el fin de proporcionar servicios. El tipo de servicios que se pueden ofrecer son: una interfaz de búsqueda, un sistema de evaluación por pares, etc. Hay que tener en cuenta que una organización proveedora puede desempeñar ambas funciones, tanto la de proporcionar datos para su recolección como la de proporcionar servicios para el usuario final.

El uso del protocolo ha resultado exitoso dentro de la estrategia de los archivos abiertos. Se podrían citar como beneficios más conocidos el servir de sustento de un nuevo modelo de

comunicación académica, y el hacer aflorar a la superficie los recursos ocultos y la interoperabilidad a bajo coste.

OAI-PMH es uno de los mayores aciertos en el acceso a bases de datos bibliográficas (y otros datos) en los últimos años. Gracias a su sencillez y facilidad de aplicación, a lo que contribuye el hecho de que el esquema básico para los metadatos sea Dublin Core (ISO 15836:200945) se ha extendido enormemente, empujando, y recibiendo a su vez el empuje, de movimientos como el OA y dando lugar a repositorios abiertos, en el sentido de los principios de difusión de contenidos que animan este movimiento, pero también en el sentido de ser susceptibles de ser recolectados.

Los repositorios institucionales han sido los más beneficiados de la rápida expansión del protocolo y de las aplicaciones basadas en él. El retraso en la aplicación de la normativa OAI por parte de las empresas de *software* para bibliotecas ha provocado que los repositorios institucionales se hayan constituido en muchos casos en una biblioteca digital paralela al sistema de gestión bibliotecario utilizado para los materiales bibliográficos.

El gran éxito de los sistemas OAI-PMH puede verse en las cifras que ofrecen algunas aplicaciones. Según *Wayback Machine*²⁹, en junio de 2003 OAIster, creado en 2002 como un proyecto de la University of Michigan Digital Library Production Services, contaba con 1.089.937 registros de 142 instituciones; en 2009, OCLC anunció que pasaba a alojar y administrar OAIster e incluyó los registros recolectados a través de OAI en WorldCat. Las cifras en 2016 ascienden a más de 30 millones de registros procedentes de 1.500 repositorios.

El hecho de que los repositorios dispongan de protocolo OAI hace que sean visibles a través del mayor catálogo del mundo o de ser recuperables a través de los grandes buscadores como *Google* que indiza los repositorios abiertos. Para ello es preciso registrar el repositorio en

²⁹ *Wayback Machine*, <https://web.archive.org/index.jsp>

diferentes registros, OAI Registered Data Providers o en los proveedores de servicios como OAlster, *Registry of Open Acces Repositories* (ROAR), Recolecta o Hispana.

Además, Hispana, <http://hispana.mcu.es/>, es el agregador español de Europeana, <http://www.europeana.eu/>, la biblioteca digital europea, y Recolecta es el agregador de DRIVER, iniciativa europea (*Digital Repository Infrastructure Vision for European Research*) que en la actualidad está integrado en OpenAire, <http://www.openaire.eu>.

Tanto para participar en Europeana como para participar en OpenAire los repositorios deben cumplir unas directrices de aplicación, además del protocolo OAI-PMH, para la adecuación de estos a la metodología y normativa específica de estos dos recolectores.

1.3.4.2. Los informes COAR sobre interoperabilidad

No hay duda de que los repositorios de acceso abierto han ido en aumento y se han convertido en un importante componente de la infraestructura de la e-investigación como bien se apunta a lo largo de los estudios de la *Confederation of Open Access Repositories* (COAR, 2012, 2015; Jo & Tsuchide, 2013).

En poco tiempo los repositorios de acceso abierto y sus servicios asociados se han convertido en un componente cada vez más importante de la infraestructura de la e-investigación (*e-Research infrastructure*). El valor real de los repositorios y su potencial para ser conectados consiste en desarrollar una red de repositorios que facilite el acceso unificado a la gran cantidad de investigación abierta y a materiales relacionados de manera que las máquinas y los investigadores puedan trabajar en nuevas vías.

Para alcanzar esa interconexión se necesita de la interoperabilidad a modo de “pegamento técnico utilizado para integrar y conectar los repositorios” (Ferrerías-Fernández, 2013b).

Así pues, la creación de ese cuerpo unificado de materiales científicos va a depender totalmente de la interoperabilidad, y en concreto, de que los repositorios sigan directrices consistentes, protocolos y estándares sobre interoperabilidad que les permitan comunicarse entre ellos, conectar con otros sistemas, y transferir información, metadatos, y objetos digitales entre sí. La juventud de la infraestructura de los repositorios produce un paisaje de interoperabilidad en evolución que a primera vista puede parecer caótico, confuso y complejo.

Como consecuencia del desarrollo alcanzado por el acceso abierto y de la mayor implicación de gobiernos e instituciones en el desarrollo de políticas que favorezcan este movimiento, se ha producido un aumento de los materiales de investigación en acceso abierto a través de repositorios y revistas.

Este aumento del conocimiento disponible libremente y de los contenedores que almacenan la información han dado lugar a un aumento de los avances en interoperabilidad de los repositorios, todos ellos diseñados para hacer frente a diversas cuestiones derivadas de este nuevo entorno, complejo y descentralizado.

Algunos de los desafíos y cuestiones planteados en este nuevo entorno conciernen al acceso y uso a la información, a los permisos y derechos, a los múltiples depósitos de un mismo artículo, a la identificación de los autores y al intercambio de información con otros sistemas de gestión de la información científica.

El informe COAR de 2012

La Confederación de Repositorios de Acceso Abierto (COAR) publicó el 26 de octubre de 2012 un informe sobre el estado actual de la interoperabilidad de los repositorios de acceso abierto bajo el título *The Current State of Open Access Repository Interoperability (2012)*. El informe ofrece una visión general del panorama actual de la interoperabilidad desde el punto de vista de los servicios que son posibles hoy en día.

El informe de COAR (2012) pretende servir de guía y perfilar los servicios y los retos de interoperabilidad que se pueden abordar. El informe proporciona una guía útil para comenzar a trabajar y señala qué iniciativas de interoperabilidad son necesarias poner en práctica para lograr determinados servicios, información que es pertinente para las instituciones que se inician en el acceso abierto y en los repositorios. Para las instituciones y gestores de repositorios que ya están involucrados en el acceso abierto y en los repositorios, el informe puede proporcionar ideas sobre funcionalidades para agregar a su repositorio o sobre servicios adicionales que pueden prestar a su comunidad.

El informe se encuadra dentro del proyecto de interoperabilidad llevado a cabo por COAR (*COAR Interoperability Project*), de acuerdo al objetivo clave que COAR marcó para el periodo 2011-2013 y que no es otro que el de “facilitar la discusión sobre la interoperabilidad entre los repositorios de acceso abierto y como parte de una más amplia infraestructura de e-investigación”. Este proyecto se ha llevado a cabo en tres etapas:

- La primera fase consistió en la puesta en marcha del proceso mediante la redacción y difusión de un breve documento con el fin de proporcionar una visión general sobre la interoperabilidad de los repositorios (2º trimestre de 2011). El resultado fue un documento informativo publicado el 11 de julio de 2011, bajo el título “The Case for Interoperability for Open Access Repositories” y del que hay una traducción al español (COAR. Grupo de trabajo 2: Interoperabilidad del Repositorio, 2011).
- La segunda fase consistió en realizar un debate abierto para reunir información acerca de los desafíos actuales y futuros relativos a la interoperabilidad global de repositorios de acceso abierto y llevar a cabo una evaluación sobre el panorama actual de interoperabilidad (3er trimestre de 2011).

- En la tercera fase, basándose en la información recopilada durante la segunda fase, el objetivo fue escribir y publicar el informe sobre “El estado actual de la interoperabilidad de los repositorios de acceso abierto” (COAR, 2012).

El informe está estructurado en dos secciones y recoge siete áreas relacionadas con la interoperabilidad, proporcionando además una introducción general sobre diecinueve iniciativas clave asociadas a esas áreas.

La sección 1 proporciona una introducción y antecedentes con el fin de mostrar la panorámica general de la interoperabilidad, la identificación de los principales retos, el estado actual y las áreas emergentes.

Las áreas recogidas en esta parte incluyen:

- Recolección de metadatos.
- Uso de estadísticas.
- Sistemas de transferencia de contenidos.
- Creación de identificadores únicos para autores y contenido de repositorios.
- Aumento de la visibilidad del contenido almacenado en repositorios a través de motores de búsqueda de Internet.
- Servicios de soporte para redes de repositorios.
- Integración de repositorios en los flujos de trabajo de la administración de la investigación.
- Soporte para los objetos compuestos (objetos digitales que incluyen componentes múltiples o archivos).

Otras áreas que se encuentran actualmente en fase de desarrollo y que no se tratan en este informe son los servicios para soportar *Linked data*, <http://linkeddata.org/>, la *Web Semántica*, <https://www.w3.org/standards/semanticweb/>, y la preservación digital.

La Sección 2 del informe proporciona la descripción de las actuales iniciativas de interoperabilidad y presenta la siguiente estructura:

Descripción – Información acerca de las directrices de interoperabilidad y estándares específicos mencionados en la hoja de ruta. Cada ítem resumido de una manera consistente.

- Aplicaciones: ejemplos concretos de servicios o aplicaciones que han sido desarrolladas utilizando estándares/directrices/proyectos.
- Recursos adicionales para otras lecturas, centrándose especialmente en los textos explicativos y otra información útil.

Las iniciativas referidas incluyen: AuthorClaim, CRIS-OAR, DataCite, DINI Certificate for Document and Publication Services, DOI, DRIVER, Handle System, KE Usage Statistics Guidelines, OAI-ORE, OAI-PMH, OA-Statistik, OA Repository Junction, OpenAIRE, ORCID, PersID, PIRUS, SURE, SWORD, y UK RepositoryNet+.

Este informe se diseñó para ser la primera etapa de un proceso multifase con el objetivo de establecer la hoja de ruta de COAR para la interoperabilidad. La segunda fase se completó con la publicación de un informe de seguimiento sobre las directrices para el futuro de interoperabilidad.

El informe de COAR aporta, de forma clara, concisa y breve, información actualizada del estado de la cuestión. Es una guía útil para ponerse al día sobre el tema de la interoperabilidad y la infraestructura que la hace posible.

Los repositorios se han convertido en componentes de infraestructura bien establecidos y han adoptado una variedad de roles en el entorno de la comunicación académica, incluida la evaluación de la investigación, el acceso abierto, la publicación y la preservación. En cada una

de estas funciones, los repositorios interactúan entre sí y con otros sistemas, lo que requiere un cierto nivel de interoperabilidad entre sistemas.

"El valor real de los repositorios es su potencial para estar conectados con el fin de desarrollar una red de repositorios que permita un acceso unificado a una masa abierta, agregada de materiales académicos y relacionados de forma que las máquinas y los investigadores puedan trabajar con ellos por vías diferentes" (COAR, 2012).

De acuerdo a las áreas de interoperabilidad descritas en el informe COAR, el repositorio Gredos tiene implementadas muchas de las herramientas que lo convierte en un repositorio interoperable (Figura 16).

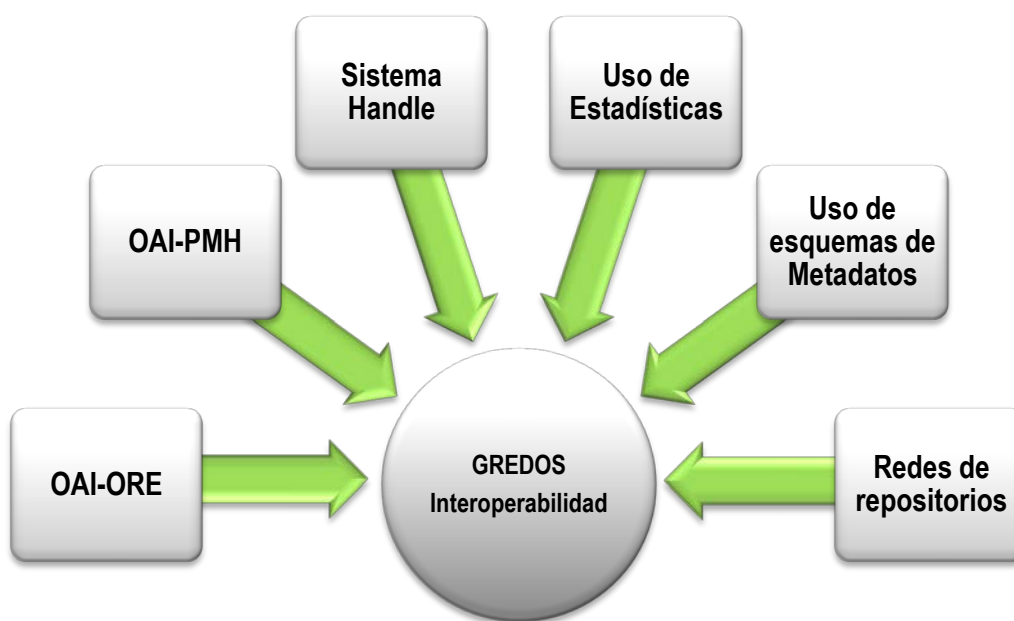


Figura 16. Áreas de interoperabilidad del repositorio Gredos. Fuente: Elaboración propia

El informe COAR de 2015

En febrero de 2015, COAR publica un mapa de ruta con las nuevas direcciones para conseguir la interoperabilidad de los repositorios (COAR, 2015).

Fruto de los rápidos avances en las tecnologías de la información y de la comunicación, el sistema de comunicación científica está experimentando un cambio fundamental hacia nuevos métodos, servicios y herramientas que den soporte al concepto de ciencia abierta.

La ciencia abierta requiere acceso transparente, uso, reutilización y confianza en la validez de los resultados de la investigación (que no solo incluyen las publicaciones y datos de la investigación, sino también las metodologías, software y hardware). La contextualización de los resultados de investigación y la vinculación de publicaciones, los conjuntos de datos y la información de los proyectos de investigación son aspectos fundamentales de este nuevo entorno.

Además, las nuevas formas de *open peer review* y las nuevas métricas del impacto de la investigación están evolucionando rápidamente. Las consecuencias de estas transformaciones aún no están claras, pero, sin duda, van a tener un fuerte impacto en los repositorios y en su función.

Según el informe de COAR (2015) un posible futuro escenario para los repositorios es que en gran medida lleguen a ser invisibles para el usuario, pero actuarán en un segundo plano como apoyo a los proveedores de servicios externos y facilitarán la visibilidad del contenido a través de los robots científicos de búsqueda, los agregadores y los indexadores. Por ello, la calidad del contenido, los metadatos, y la fiabilidad de los servicios de valor añadido serán requisitos clave para satisfacer las necesidades de los usuarios finales y para continuar siendo competitivos con servicios similares.

Esta invisibilidad conlleva el peligro inherente de que los repositorios pasen desapercibidos y sean absorbidos por la infraestructura de otros elementos, acabando con el tiempo integrándose en otros servicios.

Para que los repositorios sigan siendo relevantes en este entorno de cambio constante y rápido, la comunidad de repositorios debería adoptar una perspectiva de respuesta, de adaptación y centrarse en el desarrollo de servicios de valor para la comunidad investigadora y otros usuarios.

Las distancias cortas para el científico y una estrecha relación con la institución académica en combinación con la orientación del servicio, son ejemplos de las ventajas de jugar en casa, que los repositorios podrían utilizar para mantener y reforzar su posición. Los repositorios deberían mantener también un balance de planificación para hacer frente a los problemas prácticos de apoyo presente y para mejorar el actual nivel de calidad.

El objetivo de la hoja de ruta de COAR (2015) es definir las piedras angulares de la interoperabilidad de los repositorios en función de su relevancia y nivel de complejidad, con especial atención a los siguientes retos:

1. **Técnicos:** implementaciones de APIs, formatos de metadatos, servicios de valor añadido, vocabularios; enlaces con otras entidades (publicaciones, datos de investigación, información de proyectos, estadísticas de impacto; intercambio de activos digitales; cobertura y soporte de activos digitales más allá del texto.
2. **Organizativos:** funciones y responsabilidades de funcionamiento, soporte, desarrollo.
3. **Legales:** cuestiones sobre el intercambio de datos y la reutilización.

Las diversas funciones de los repositorios y sus relaciones con otros sistemas existentes o en evolución abogan por la necesidad de interoperabilidad entre numerosas interfaces e identidades.

Sin embargo, no está claro en cuál de las múltiples áreas de interoperabilidad debe centrar sus esfuerzos la comunidad del repositorio. A tal fin, el Grupo de Trabajo de Interoperabilidad de COAR puso en marcha un proceso para ayudar en la identificación de prioridades en el trabajo sobre interoperabilidad, y para aumentar la conciencia de los problemas de interoperabilidad en

una comunidad más amplia. Desarrollaron una lista completa de 6 tópicos de interoperabilidad, dividido en 35 temas, como se muestra en la Tabla 30. Estos tópicos se clasificaron para su implementación teniendo en cuenta la necesidad estratégica, el nivel de relevancia y el grado de complejidad para su implementación.

Tabla 30

Tópicos y temas de interoperabilidad

<p>Impacto y Visibilidad</p> <ul style="list-style-type: none"> • Soporte para <i>Search Engine Optimization</i> (SEO) • Soporte para sistemas de <i>Ranking</i> de Repositorios • Exposición de Estadísticas de uso • Exposición de información bibliométrica • Soporte de visibilidad en los registros del repositorio • Mejora de la infraestructura de registros <p>Usabilidad</p> <ul style="list-style-type: none"> • Soporte para autorización y autenticación • Soporte para servicios de incrustación • Exposición de listas de publicaciones • Exposición de formatos de citas • Soporte para funciones de exportación de datos • Integración de servicios de disponibilidad • Soporte para sistemas de identificación de autor • Soporte para servicios institucionales • Ampliación de la usabilidad para el usuario final • Ampliación del uso de herramientas de visualización <p>Sostenibilidad</p> <ul style="list-style-type: none"> • Mejora de la estabilidad de la plataforma • Soporte para la preservación a largo plazo y para el archivo • Exposición de identificadores persistentes • Integración de diferentes identificadores persistentes 	<p>Cuestiones sobre datos</p> <ul style="list-style-type: none"> • Soporte para formatos de metadatos adicionales • Mejora de la calidad de los metadatos (curación de datos) • Soporte para publicaciones mejoradas • Soporte para <i>Linked Open Data</i> • Publicación de datos de investigación • Manejo de Objetos complejos, compuestos y anidados del repositorio • Monitorización del cumplimiento del mandato de acceso abierto • Exposición de la información de las versiones <p>Validación y agregación</p> <ul style="list-style-type: none"> • Validación de metadatos del repositorio • Procesamiento relacionado con el texto completo • Sobre duplicación <p>Cuestiones técnicas</p> <ul style="list-style-type: none"> • Definición de recomendaciones sobre la arquitectura de los repositorios y de su interoperabilidad • Ampliación / sustitución de los protocolos de exposición de metadatos • Soporte <i>OAI Service Provider Usage</i> • Soporte de protocolos de depósito
---	--

Nota: Basado en COAR (2015)

Los resultados de la hoja de ruta demuestran que el intercambio y la reutilización de los metadatos de los repositorios representan una de las principales prioridades para el trabajo

sobre interoperabilidad. Esto servirá de apoyo para la interoperabilidad entre repositorios y sistemas (por ejemplo, con CRIS y sistemas biográficos de investigación). También trata la crucial importancia que tiene para la comunidad repositorio la integración con otros sistemas a fin de proporcionar información estandarizada a efectos de los sistemas de administración de la investigación. Para lograr esto, también serán clave los vocabularios estandarizados y los esquemas de metadatos y de elementos así como los sistemas de identificación unívoca de los investigadores.

I.3.5. Difusión, visibilidad, uso e impacto de los contenidos de los repositorios

El acceso abierto aumenta las posibilidades de difusión de una publicación, ya que no se limita su consulta desde una única fuente (revista, congreso o editorial, por ejemplo), sino que al estar también depositado en un repositorio abierto, los distintos sistemas de recuperación y recolección permitirán que los datos de ese documento estén accesibles desde múltiples fuentes, lo que amplía la difusión y visibilidad de los documentos de los repositorios abiertos. El acceso abierto es una opción excepcional para la difusión de los resultados de la investigación entre la comunidad científica internacional.

La Web, las bibliotecas digitales y los repositorios han cambiado la forma de difundir el conocimiento científico, ofreciendo nuevas oportunidades para obtener una mayor y exhaustiva cobertura de la literatura científica (Aguillo et al., 2010).

Las revistas de acceso abierto y el depósito de los resultados de investigación en repositorios no solo se ven favorablemente afectados por la visibilidad (Aguillo et al., 2010; Antelman, 2004; Eysenbach, 2006; Hajjem et al., 2005; S. Harnad & Brody, 2004; Kurtz et al., 2005; Moed, 2007; Norris, Oppenheim, & Rowland, 2008) sino que abren la evaluación de la investigación a metodologías más centradas en el usuario.

Los repositorios mantienen la promesa de contribuir de manera significativa a las tasas de citación del autor. Los repositorios institucionales pueden desempeñar un papel crucial en la medición de resultados de la investigación, y a su vez pueden afectar a los *ranking* universitarios. En este sentido, el *Ranking Web de Repositorios*³⁰ pretende motivar, tanto a instituciones como a investigadores, a tener una presencia en la web que refleje fielmente sus actividades. Afirman que si el rendimiento web de una institución está por debajo de lo esperado de acuerdo a su excelencia académica, las autoridades de la institución deberían reconsiderar su política Web, promoviendo el incremento substancial del volumen y la calidad de sus publicaciones electrónicas abiertas.

Algunos investigadores realizaron un estudio sobre una base de datos de más de un millón de artículos publicados en el período de 1992-2003 en diez disciplinas científicas, y llegaron a la conclusión de que los documentos en acceso abierto han recibido constantemente más citas que los que no están en acceso abierto (Hajjem et al., 2005). Otros autores encontraron que la proporción anual de documentos de acceso abierto ha estado creciendo mucho más rápido que la de sus homólogos que no están en abierto, mediante la comparación de los dos tipos de artículos en la misma revista y año (S. Harnad & Brody, 2004).

Se constata pues a través de los estudios de estos autores que el acceso abierto contribuye al crecimiento de las citas de un documento. La filosofía en la que se basa la creación de repositorios institucionales, así como las tecnologías que emplean, posibilitan el aumento del factor de impacto de las publicaciones y, por tanto, de sus indicios de calidad, al ofrecer públicamente artículos científicos y cualquier documento proveniente de una investigación. Difundir publicaciones de forma pública, a través de repositorios accesibles de manera abierta,

³⁰ "Objectives of the Ranking Web of World's Repositories", available at: <http://repositories.webometrics.info/en/Objectives> (accessed 30 May 2014).

permite que el alcance de la producción científica sea universal y no se limite a circuitos comerciales cerrados como es el de la publicación en revistas científicas.

En los repositorios se permite que los autores depositen diferentes versiones de sus trabajos científicos: obras inéditas y publicadas, *pre-prints* y artículos a examen, actualizaciones y presentaciones, documentos textuales o multimedia, etc. El circuito de la publicación científica se ve beneficiado por el acceso abierto al disponer de mecanismos rápidos para la exposición de los resultados de una investigación, para difundir una investigación publicada, para actualizar contenidos o para actuar como masa crítica.

El acceso abierto también facilita la creación de redes científicas, al posibilitar el conocimiento mutuo de autores que trabajan en los mismos ámbitos, pudiendo ser un excelente punto de partida para la creación de grupos de investigación y para el seguimiento de la producción de determinados autores, que publicarán en diferentes fuentes, pero que alojarán sus trabajos en sus correspondientes repositorios institucionales; aunque seguir a un autor en su repositorio se hace complicado por la poca experiencia de usuario que presentan los repositorios actuales y que les convierte en un buen almacén que interopere con otros sistemas, pero de escasa utilidad para la mayoría de los usuarios humanos (excepto expertos bibliotecarios). Por ello, un paso necesario será el desarrollo y evolución de los *discoveries* (Breeding, 2015).

De igual forma, el depósito de producción científica en acceso abierto es útil para el conocimiento de trabajos interdisciplinarios y de aquellos que puedan interesar a diferentes comunidades de científicos. Las investigaciones especializadas y las multidisciplinares encuentran en el acceso abierto un excelente medio para ser difundidas y alcanzar una masa crítica superior a la que ofrecen los canales tradicionales de publicación.

Las revistas en acceso abierto que publica una entidad, así como el volumen y calidad de los documentos depositados en su repositorio científico ofrecen indicadores cualitativos y

cuantitativos de la producción científica de una universidad o instituto de investigación. El análisis de los documentos depositados en repositorios institucionales posibilita el establecimiento de valoraciones del rendimiento científico de los organismos académicos y de investigación.

El acceso abierto aumenta las posibilidades de acceder a recursos de investigación sin aumentar el presupuesto destinado a suscripciones de colecciones electrónicas o recursos documentales para la investigación. El planteamiento del que parten los repositorios abiertos es facilitar de forma gratuita la consulta y descarga de los documentos que alojan, por lo que los investigadores pueden acceder a colecciones electrónicas de calidad a texto completo.

Capítulo I.4. Ecosistemas tecnológicos. El repositorio Gredos

I.4.1. Ecosistemas tecnológicos: una aproximación

Más allá de las modas en las tendencias tecnológicas, en las instituciones coexisten muchas colecciones de aplicaciones software, tanto comerciales como open source, que requieren de una integración e interoperabilidad para soportar su funcionamiento efectivo en el contexto corporativo (Forment et al., 2010; García-Peñalvo, 2015). Cuando el grado de integración de estos componentes es muy alto aparecen relaciones simbióticas entre ellos que obliga a atender tanto las necesidades de evolución de cada uno de ellos como la influencia que dicha evolución tiene en el propio contexto corporativo, incluyendo a los propios usuarios como otros componentes más, surgiendo así un ecosistema tecnológico de alta complejidad (García-Holgado & García-Peñalvo, 2013b, 2016).

El concepto “ecosistema tecnológico” es una metáfora utilizada para ejemplificar la manera en la que funcionan y se integran los diferentes elementos y herramientas que forman parte de la estrategia tecnológica de una organización. Se denomina ecosistema porque es comparado con el ecosistema biológico, como comunidad de seres vivos cuyos procesos vitales están interrelacionados y cuyo desarrollo se basa en los factores físicos del medio ambiente. Cuando se intenta trasladar esta acepción biológica al contexto tecnológico existen múltiples definiciones, con distintos puntos de vista, pero todos ellos confluyen en un punto fundamental, hay una clara relación entre las características de un ecosistema natural y un ecosistema tecnológico en cualquiera de sus variantes (García-Peñalvo, 2015).

Según Llorens Largo (2014), el presente está protagonizado por la integración, que conecta y relaciona las distintas herramientas que van surgiendo y que sirven para la labor docente, construyendo ecosistemas tecnológicos. Pero esas herramientas integradas en ecosistemas no

solo sirven para la labor docente, sino que su utilidad se extiende a todos los ámbitos de la institución llamada Universidad (investigación, docencia, estudio y servicios). En definitiva, se está hablando de Universidad Digital.

Para Chang y West (2006) los ecosistemas digitales trascienden los tradicionales entornos colaborativos desde modelos centralizados, distribuidos o híbridos hacia un entorno abierto, flexible, basado en la demanda e interactivo. Un ecosistema digital es una nueva arquitectura de trabajo en red y entorno colaborativo que se dirige a los puntos débiles de los servicios cliente-servidor, *peer-to-peer*, *Grid* y *web*.

Actualmente, la gestión del conocimiento es una de las principales necesidades que tiene cualquier tipo de organización o institución. La gestión del conocimiento está asociada no solo con la gestión del conocimiento como recurso sino también con la gestión de los procesos que tienen lugar cuando se usa el recurso. El conocimiento por sí mismo no se puede manejar, se necesitan procesos para fomentar el intercambio de conocimientos, y el desarrollo de los activos de capital intelectual son los elementos que pueden ser manejables. La gestión del conocimiento requiere de un procedimiento que pueda ser utilizado por todos los miembros de la organización para compartir sus procesos de conocimiento y sus experiencias (García-Holgado & García-Peñalvo, 2016).

El *software* libre y los desarrollos de *Open Source* son los principales términos usados para hablar de herramientas tecnológicas generadas por las empresas y las instituciones. Este tema está unido a un gran número de cambios en el contexto tecnológico. En los últimos diez años, la Web 2.0, y la evolución de los dispositivos utilizados para acceder a Internet ha llevado a un cambio de paradigma en el desarrollo de componentes de software que son a la vez propietario y de código abierto.

En este contexto, la solución a los diferentes problemas y objetivos relacionados con la gestión del conocimiento estaría basada en la definición e implementación de ecosistemas tecnológicos. Este sistema estaría compuesto por un conjunto de módulos que provean de las necesarias funcionalidades para gestionar los procesos internos y externos de la empresa y para soportar los flujos de información entre los diferentes componentes del sistema (García-Holgado & García-Peñalvo, 2013b).

Los ecosistemas tecnológicos podrán ser considerados como un marco general para desarrollar cualquier tipo de solución tecnológica donde la información es el centro del sistema. El ecosistema puede estar orientado hacia diferentes áreas dependiendo de los problemas que deban resolverse. Existen ecosistemas tecnológicos para propuestas generales que pueden ser fácilmente extendidos y desarrollados para propuestas específicas. En el caso de la gestión del conocimiento orientada a proporcionar y mejorar la formación dentro de una institución, el ecosistema tecnológico se llama el ecosistema eLearning (F. J. García-Peñalvo et al., 2011). Este tipo de ecosistema está adaptado a la evolución de la gestión del conocimiento que se lleva a cabo en empresas e instituciones.

El desarrollo de los ecosistemas tecnológicos es diferente en las empresas e instituciones. A pesar de que los principales objetivos son los mismos, los componentes de software y flujos de información pueden cambiar incluso en la misma entidad. Dentro del Grupo de Investigación en la interacción y aprendizaje electrónico (GRIAL) de la Universidad de Salamanca (García Peñalvo et al., 2012), los autores han participado en el desarrollo de varios ecosistemas tecnológicos orientados a la gestión del conocimiento en diferentes contextos para resolver problemas reales (García-Holgado & García-Peñalvo, 2016; García-Peñalvo, 2008c).

En la Universidad de Salamanca se desarrolló un ecosistema tecnológico dentro de la Universidad Virtual con el fin de gestionar el conocimiento generado por la totalidad de la

comunidad universitaria (García-Peñalvo, 2008c; Laviña Orueta & Mengual Pavón, 2008), dentro del cual se puso en marcha el repositorio Gredos (Seoane-Pardo, 2014).

Como apunta Valverde Berrocoso (2013), dentro del marco de los ecosistemas de e-learning una solución para generar un ecosistema de este tipo es la creación de Repositorios de Objetos de Aprendizaje Reutilizables o Repositorios Institucionales, ya que son bases de datos que ofrecen servicios de captación, almacenamiento, indexación, preservación y difusión en acceso abierto de contenidos educativos y de investigación en formato digital pertenecientes a una comunidad universitaria.

I.4.2. El repositorio Gredos

El repositorio Gredos es una pieza importante del ecosistema tecnológico de la Universidad de Salamanca dentro del Proyecto Universidad Virtual. También es una pieza de la e-infraestructura de la que habla COAR (2012), y que se podría considerar un ecosistema tecnológica de interoperabilidad.

Gredos (Gestión de REpositorio DOcumental de la universidad de Salamanca, <http://gredos.usal.es>), repositorio institucional basado en DSpace, <http://www.dspace.org>, y concebido como una herramienta de procesamiento, almacenamiento y recuperación de colecciones en formato digital producidas o alojadas en la universidad. La totalidad del contenido está disponible en acceso abierto, con licencia "*Creative Commons Reconocimiento-No comercial-Sin obras derivadas 3.0 España*" (<http://creativecommons.org/licenses/by-nc-nd/3.0/es>), que preserva los derechos de autoría de los materiales almacenados.

En mayo de 2009, Gredos contaba con 272 comunidades, 2.157 colecciones, un total de 52.056 registros y 53.053 objetos digitales, e incluía más de 150.000 digitalizaciones. Según la última memoria del Servicio de Bibliotecas correspondiente al curso 2015-2016, (Universidad de

Salamanca. Servicio de Bibliotecas, 2016), las cifras alcanzan las 524 comunidades, 5.723 colecciones y un total de 111.466 registros.

Gredos se estructura en cuatro secciones:

1. La Biblioteca Digital reúne las colecciones históricas y documentos de carácter patrimonial de la Universidad de Salamanca. Esta sección recoge libros de carácter patrimonial y en dominio público de la Biblioteca Histórica de la Universidad. Además aquí se reúnen los proyectos externos de colaboración, versiones digitales de manuscritos, libros antiguos, prensa histórica, revistas españolas y otras publicaciones digitalizadas. En julio de 2016 contenía 63.277 registros. El contenido de las colecciones de esta comunidad se reparte de la siguiente forma (ver Figura 17): libros (10%), artículos (59%), imágenes y sonido (2%), y otros tipos (31%).

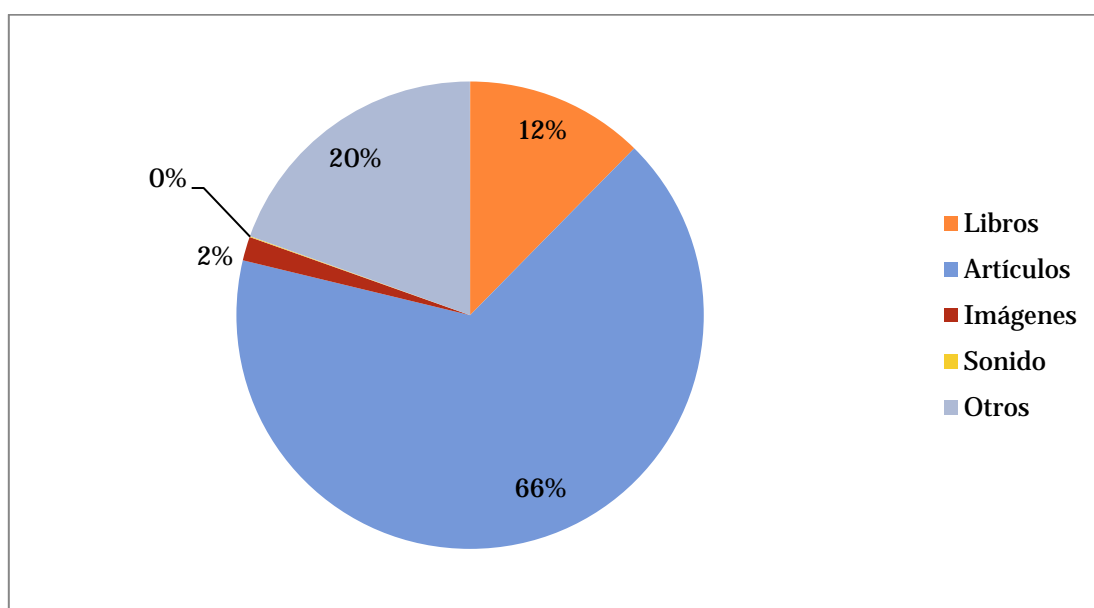


Figura 17. Contenido de la comunidad Biblioteca Digital de Gredos. Fuente: Elaboración propia

2. Repositorio científico, donde se alberga la producción científica de la Universidad de Salamanca: tesis doctorales, artículos, monografías y congresos científicos, ediciones de la universidad y revistas especializadas. Esta sección contenía 23.956 registros, en julio de 2016. El contenido de las colecciones de esta comunidad se reparte de la siguiente forma además de

artículos científicos (66%), contiene imágenes (21%), tesis doctorales (9%), monografías científicas (2%) y ponencias (1%), como se observa en la Figura 18.

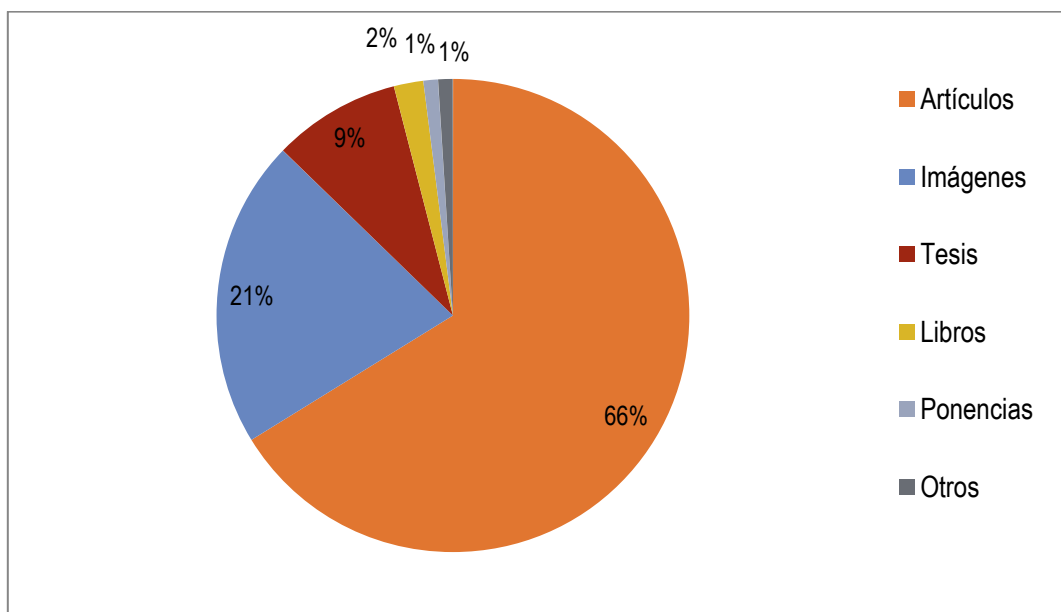


Figura 18. Contenido de la comunidad Repositorio Científico de Gredos. Fuente: Elaboración propia

Es de destacar los artículos depositado en esta comunidad de Gredos procedente de las revistas editadas por Ediciones Universidad de Salamanca (EUSAL) a través del Open Journal System (OJS). Este servicio ofrecido a través del repositorio es fruto de la colaboración del Repositorio Institucional con Ediciones Universidad de Salamanca.

3. El repositorio docente agrupa los recursos de carácter didáctico generados en la Universidad (ver Figura 19): materiales de asignaturas (14%), tutoriales o vídeos educativos (2%), así como los proyectos fin de carrera (9%), trabajos de fin de máster (23%), trabajos de grado (14%) y memorias de innovación docente (38%). por último, el Repositorio Docente ofrece Trabajos de Grado, Trabajos de Fin de Máster, Proyectos Fin de Carrera, materiales didácticos, tutoriales y asignaturas abiertas, materiales audiovisuales y otros recursos docentes. En julio de 2016 el contenido de esta sección era de 2.624 registros.

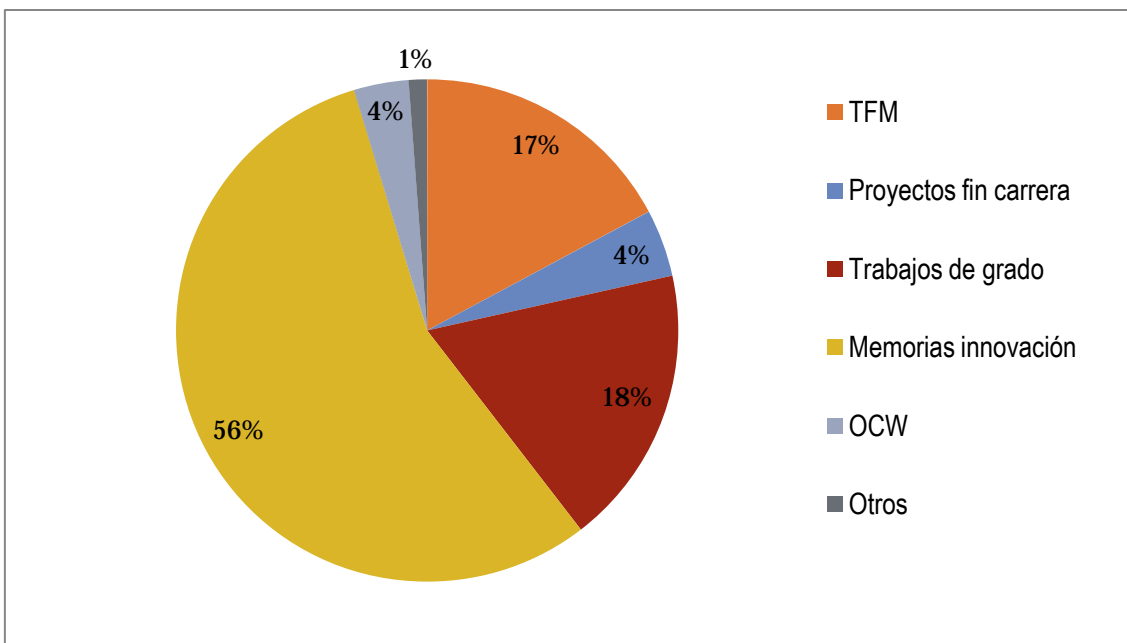


Figura 19. Contenido de la comunidad Repositorio Docente de Gredos. Fuente: Elaboración propia

4. El archivo institucional contiene documentos internos de la Universidad de Salamanca como normativas, guías académicas, partituras, dibujos, discursos o fotografías y colecciones personales de profesores de esta Universidad con todo tipo de documentación predominando las colecciones epistolares de archivo. En la figura 20 se muestra el contenido de la comunidad Archivo Institucional de Gredos. En julio de 2016 el contenido de esta sección era de 23.256 registros.

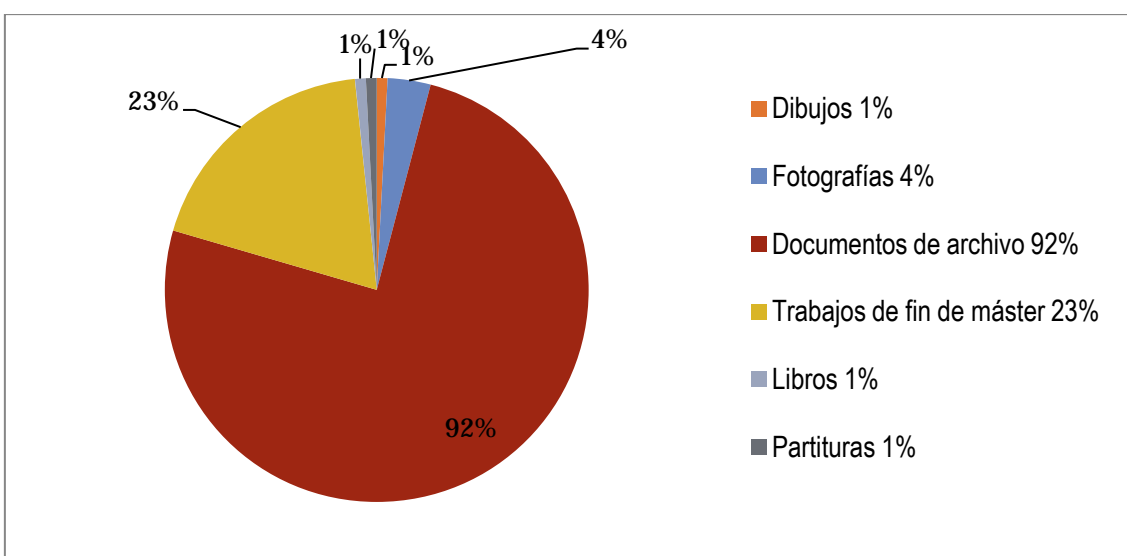


Figura 20. Contenido del Archivo Institucional de Gredos. Fuente: Elaboración propia

Gredos se puso en marcha en 2009 con más de 51.300 registros y actualmente, en agosto de 2016, supera ya los 104.000, posicionándose en el primer lugar de los repositorios institucionales españoles en cuanto a volumen de contenido se refiere, según datos del directorio ROAR. Los objetos digitales que contiene son digitalizaciones de originales en papel y otra gran parte son documentos nacidos digitales. Para gestionar toda esta información Gredos utiliza como *software* de gestión la plataforma DSpace, en su versión 1.6.2. que en julio de 2016 se encuentra en fase de actualización a la versión 5.1.

Además de recoger, almacenar y difundir información digital, se pretende que esa información perdure, que esté disponible y que sea recuperable a lo largo del tiempo sin que se vea afectada por los inevitables cambios tecnológicos que se vienen produciendo.

DSpace en Gredos se administra aplicando los protocolos y las directrices recomendadas para la preservación: modelo OAIS, METS, y metadatos de preservación (PREMIS), de esta forma se asegura la perdurabilidad a largo plazo de los objetos y contenidos digitales del repositorio y se facilita además su accesibilidad, su interoperabilidad y su consulta en el futuro.

Otro de los objetivos es asegurar la máxima difusión de sus contenidos y así conseguir que los investigadores que depositan aquí sus trabajos alcancen citación y por tanto impacto en el mundo científico. Gredos al ser recolectado por los principales recolectores nacionales -Hispana, Recolecta, TDR - e internacionales -Europeana, Oaister/OCLC, EROMM, DART - contribuye a esta difusión del conocimiento. En este sentido, es de destacar la contribución de Gredos a Europeana realizada a través de Hispana. Todo el contenido del repositorio Gredos es accesible a través de la gran biblioteca digital europea.

Además Gredos es proveedor de datos de OpenAire siendo compatible con la infraestructura requerida de OpenAIRE 2.0+ (DRIVER OA, EC funding). En la figura 21 se muestra los datos de la ficha de Gredos proporcionados por OpenAire.


GREDOS

Name	GREDOS
Type	Institutional Repository
Items	100501 Publications
Compatibility	OpenAIRE 2.0+ (DRIVER OA, EC funding)
OAI-PMH	http://gredos.usal.es/oai/request

Figura 21. Datos de Gredos proporcionados por OpenAire. Fuente: OpenAire

OpenAire recolecta también la totalidad del repositorio y ofrece datos estadísticos de las aportaciones de Gredos. En la figura 22 se muestra los datos proporcionados por OpenAire, en julio de 2016, sobre los porcentajes de los tipos de contenidos de Gredos.

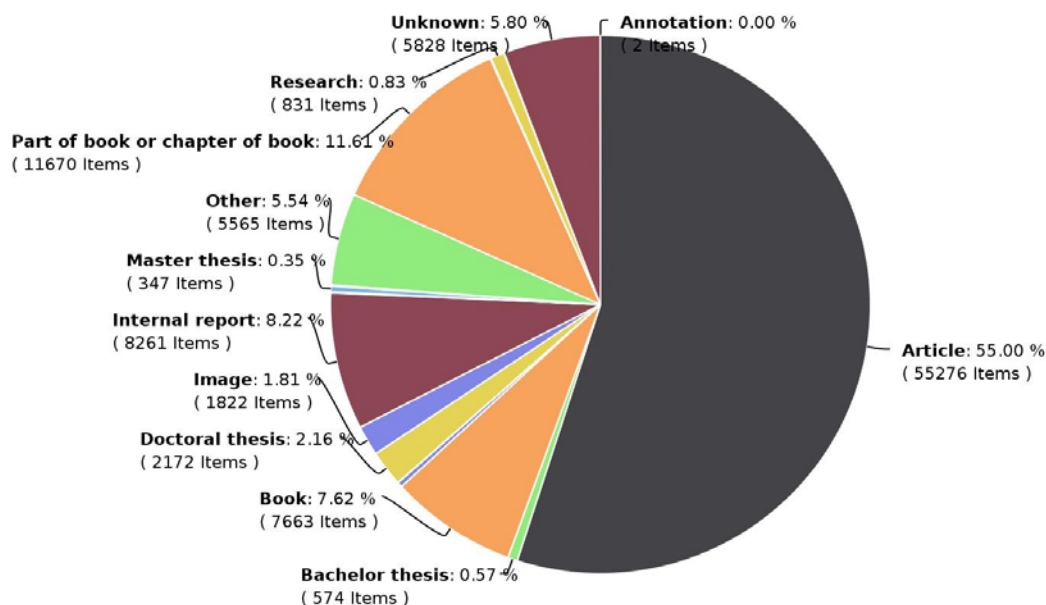


Figura 22. Tipos de contenidos de Gredos. Datos extraídos de OpenAire. Fuente: OpenAire (julio 2016)

Por otra parte es de destacar la presencia de Gredos en *Google Scholar* que al indizar los pdf completos del repositorio es el portal que le proporciona mayor visibilidad en razón de los datos aportados por las estadísticas de descargas y accesos.

A fecha de septiembre de 2016 está en marcha una actualización del repositorio Gredos que comprende la actualización de la versión de Dspace, un nuevo diseño de la interfaz, la implementación de herramientas altmétricas (Plum Analytics, Almetric, Google Scholar).

Capítulo I.5. La literatura gris científica. Las tesis doctorales

En este capítulo se ofrece una panorámica sobre la definición y evolución de la Literatura Gris (LG) en el entorno digital y del acceso abierto a la información científica y técnica. Se hace especial hincapié en las tesis doctorales que son el foco central de la investigación de esta tesis. Se trata el tópico del aumento de las e-tesis en acceso abierto y los beneficios, cambios y compromisos que esta nueva forma de comunicación reporta a las e-tesis.

I.5.1. Literatura gris. Definición y datos

La literatura gris está omnipresente. Según las disciplinas, la LG representa hasta el 30% de la producción científica. Los análisis cuantitativos muestran su impacto a nivel de citas, además la LG constituye una parte importantes de colecciones, portales y archivos abiertos (Schöpfel, 2012).

Sin embargo el término “literatura gris” a veces se define mal, con imprecisiones, y límites débiles. El concepto es histórico aunque en algún tiempo no se conocía la literatura gris. El concepto se propagó a partir de los años 80, aunque había ido emergiendo gradualmente desde los años 70. Cuando Charles P. Auger (1975) publicó la primera edición de su obra sobre literatura de informes en 1975, no usó el término “literatura gris”. No obstante las referencias a las características de la literatura gris son reconocibles en esta obra. Auger hablaba en un primer momento de informes técnicos y científicos. Otros autores intentaron precisar este concepto y extenderlo como término genérico a otros documentos que tenían las mismas características, sobre todo las tesis pero también las comunicaciones, los documentos de trabajo, manuscritos (*pre-prints*), etc. *Don't be afraid of reports* es el título de un artículo escrito por Chillag (1973) que resume la tendencia recurrente de un período que ahora parece anticuado.

El desarrollo de las redes ha reducido muchos de los problemas relacionados con la recuperación y difusión de información, ayudando a producir diferentes formas de literatura gris y diferentes métodos de recolección y manejo de este tipo de documentación (Luzi, 2000).

Hay muchas definiciones de literatura gris, la más conocida es la llamada "definición de Luxemburgo", que fue discutida y aprobada durante la 3ª conferencia Internacional sobre Literatura Gris (1997): "Literatura gris es aquella se produce a instancias del gobierno, de la enseñanza e investigación pública, del comercio y de la industria en formato impreso o electrónico, pero que no está controlada por editores comerciales" ("Third International Conference on Grey Literature : Perspectives on the Design and Transfer of Scientific and Technical Information, 13-14 November 1997," 1997) . Esta definición en sí contiene dos de las características principales de los recursos "grises": por una parte son universales y ubicuos, pero por otra parte, son difíciles de identificar y de obtener a través de los circuitos tradicionales de publicación. La definición de Luxemburgo es también bastante vaga para reflejar el problema de cómo definir exactamente qué es lo que significa realmente un tipo de literatura descrita a veces como "subterránea", "efímera" o "no convencional" (Luzi, 2000).

El U.S. Interagency Gray Literature Working Group precisó que se trata de documentos que no entran en los sistemas habituales de edición, de distribución, de control bibliográfico o adquisición a través de bibliotecas o agencias de suscripción mientras que Mackenzie Owen (1998) observa que "gris no implica ninguna calificación, sino que es simplemente una característica del modo de distribución".

Algunos años más tarde, en 2004, en la sexta conferencia sobre literatura gris de Nueva York se precisó la definición de Luxemburgo incluyendo una posdata: "literatura gris es la literatura que no está controlada por los editores comerciales". Desde entonces esta definición fue la más

extendida y utilizada ("Sixth International Conference on Grey Literature: Work on Grey in Progress, 6-7 December 2004," 2004).

Alrededor del 40% de los estudios adoptan una aproximación tipológica. La cuestión clave es saber qué documentos pertenecen a la LG. Los bibliotecarios están de acuerdo en que las tesis, las actas de conferencias, los informes y los documentos de trabajo son LG. Pero también hay otro tipo de documentos como patentes, prepublicaciones, blogs, bases de datos, etc. que se pueden clasificar como LG.

El término tradicionalmente abarcaba tres categorías de documentos: conferencias, informes y tesis doctorales, que frecuentemente se imprimían en tiradas cortas. Sin embargo, el límite con literatura "blanca" o "convencional" es permeable desde que algunas actas de conferencias se publican por editores comerciales, como monografías en publicaciones seriadas o revistas, y lo mismo ocurre con algunos informes. En cuanto a las tesis doctorales, sobre todo en las áreas de humanidades y ciencias sociales algunas se pueden encontrar en el mercado de la edición comercial.

De acuerdo con Luzi (2000), el mayor avance ha sido el canal de difusión que, antes de la llegada de las redes, marcaba una de las diferencias sustanciales entre la literatura convencional y la no convencional. El cambio, acompañado de un fuerte aumento de la producción de la información en red, ha contribuido a modificar la naturaleza misma de la LG. La dificultad en la recuperación, que solía ser un rasgo distintivo de la LG, sobre la que se centraron la mayor parte de los estudios antes de la difusión a gran escala por Internet, se ha ido reduciendo. Al mismo tiempo, el uso de Internet como un canal de difusión tanto de la literatura convencional y como de la no convencional también ha dado lugar a un cambio de perspectiva para muchos estudios, que se centran en cuestiones que ya no son solo peculiares a la LG sino comunes a gran parte de la información disponible en la Web. Cuestiones tales como la actualización y validación de la

información o los derechos de autor se convierten así en los temas de actualidad de discusión, dentro de un marco más amplio de la gestión y difusión del conocimiento.

Actualmente, la LG está universalmente reconocida como un importante recurso de información. Contiene mucho conocimiento para las organizaciones industriales y comerciales. La literatura gris en un entorno de I + D puede representar la vanguardia del conocimiento, por lo que su gestión eficiente es esencial. Las posibilidades, cada vez mayores, que trae consigo la publicación científica electrónica junto con el desarrollo de la Web, son las fuerzas impulsoras de la implementación de proyectos para el archivo, la preservación y la provisión de acceso a las colecciones de literatura gris científica electrónica (Correia & Neto, 2002).

La literatura gris es una importante fuente de información como lo es la literatura convencional, no diferenciándose entre sí en valor científico sino en su contenido y finalidad, así como en los métodos de producción y difusión. Por otra parte, con el uso de Internet se ha eliminado casi por completo uno de los problemas que afectaban a la literatura gris: su disponibilidad.

El acceso abierto está cambiando gradualmente la concepción de la comunidad científica sobre los documentos que no están sujetos a revisión por pares (y por tanto excluidos de la evaluación a efectos de carrera), en el proceso de acercar la literatura gris a la literatura blanca. Además, la literatura gris siempre ha sido parte de la red de comunicación científica como se demuestra por las citas de documentos de literatura gris en las referencias bibliográficas de artículos científicos. Como indican Cesare, Luzi y Ruggieri (2007) hasta cierto punto, la literatura gris es una forma *ante litteram* de la comunicación académica. Las autoras indican que a la luz de estos acontecimientos, el foco de atención en LG se ha desplazado desde las dificultades de definición y tipología, que son casi irrelevantes para los usuarios, hacia el análisis de las referencias bibliográficas, que constituyen la base común entre la literatura convencional y la no convencional. En su artículo, las autoras proponen una metodología para analizar el uso de LG

en las referencias bibliográficas de artículos de revistas, que ya se aplican en algunos campos disciplinarios.

La literatura gris tiene un papel propio como medio de difusión de información científica y técnica (Schöpfel, 2006; Søndergaard, Andersen, & Hjørland, 2003), y los profesionales insisten en su importancia por dos razones principales: los resultados de la investigación son a menudo más detallados en los informes, tesis doctorales y actas de la conferencia que en las revistas y se distribuyen en estos formatos hasta 12 o incluso 18 meses antes de ser publicados en otro lugar.

I.5.2. Desarrollo de políticas para la gestión de Literatura Gris

En abril de 2014, profesionales de la información europeos se reunieron en Pisa para abordar temas relacionados con la política y la gestión de los recursos de literatura gris. Esta reunión dio lugar a la formulación de la Declaración de Pisa sobre el Desarrollo de Políticas de Recursos de Literatura gris, una hoja de ruta de 15 puntos que sirve como guía para las organizaciones que participan en la producción, publicación, el acceso y uso de la literatura gris en el siglo XXI. La Declaración de Pisa, al igual que las declaraciones de Budapest, Bethesda y Berlín ofrece un enfoque diferente en el tratamiento de la información y de los datos con el fin de satisfacer las necesidades y requerimientos del mercado hoy en día. Hasta ahora, el problema era la falta de cooperación y coordinación entre y dentro de las organizaciones que se ocupan de la literatura gris. La Declaración de Pisa, <http://greyguide.isti.cnr.it/>, señala el fin de una era de políticas *ad hoc* y de toma de decisiones con respecto a recursos de LG. Desde su publicación original en inglés, la Declaración de Pisa ha sido traducida a 20 idiomas, <http://greyguiderep.isti.cnr.it/Pisadeclapdf/Spanish-Pisa-Declaration.pdf>, entre ellos el español, y se prevén más traducciones próximamente.

En la introducción de la Declaración de Pisa se ofrece también una definición de literatura gris: existe un caudal de información y de datos que producen las organizaciones, los gobiernos y la

industria, que abarca una gran cantidad de áreas temáticas y campos profesionales, esos datos y otros materiales derivados, se conocen como literatura gris y son un recurso esencial en la comunicación académica, en la investigación y en el diseño de políticas para las empresas, la industria, la práctica profesional y la sociedad civil. Además, se reconoce la literatura gris como una fuente primordial de evidencia, discusión, innovación y comprensión en muchas disciplinas, tales como las ciencias, las ingenierías, la salud, las ciencias sociales, la educación, las artes y las humanidades.

También se definen los tipos documentales de la LG que se producen en formatos impresos o digitales, incluyendo entre otros los informes técnicos y de investigación, las ponencias, las tesis, y los contenidos en multimedia. En la misma Declaración se hace un llamamiento para que se dé un mayor reconocimiento del valor de la LG desde los gobiernos, las universidades y todos los grupos involucrados con el fin de que los beneficios en investigación y en información sean patentes para la ciencia, los gobiernos, la educación, la economía y en general para toda la sociedad. Asimismo, y este punto se relaciona directamente con el objeto de esta tesis, en la Declaración de Pisa se enfatiza la importancia de la literatura gris para la investigación de acceso abierto, para una ciencia abierta, que involucre la innovación, el diseño de políticas basado en la evidencia y en la transferencia de conocimiento.

I.5.3. La evaluación de Literatura Gris

Tanto los individuos como las organizaciones e instituciones que producen conocimiento para los demás necesitan ser capaces de proporcionar evidencias del valor de su trabajo, de la misma forma que los científicos pueden utilizar el factor de impacto de las revistas y las citas de sus trabajos para indicar el valor de sus artículos (Wilkinson, Sud, & Thelwall, 2014).

Sobre la evaluación de la literatura gris existen ya algunos estudios y cada vez son más los autores que se interesan por encontrar un sistema de medición del impacto de LG que demuestre el valor de este tipo de literatura como fuente importante de información.

Di Cesare y Ruggieri (2006) propusieron una metodología para analizar el uso de LG en referencias bibliográficas de artículos de revistas, aplicados en diferentes disciplinas. El método de medición utilizado tenía una base teórica de los conceptos e instrumentos utilizados en los estudios de citación (Garfield, 1970). Independientemente del hecho de que no exista ninguna teoría genuina de citación (Leydesdorff, 1998), y que la aplicación de los instrumentos bibliográficos no llegan a conclusiones válidas en términos generales, no deja de ser una herramienta importante para medir y comparar la literatura científica y técnica.

Diferentes estudios han utilizado instrumentos bibliométricos para evaluar el impacto de LG en la literatura convencional. Algunos de estos instrumentos se utilizan para evaluar el impacto de LG de un tipo particular de documento y su productor (Cordes, 2004) en un campo o disciplina especial, y rastrear la existencia o no de la interdependencia entre la literatura convencional y literatura gris (Dunn, 2005). El método ha sido utilizado por otros para seguir el trabajo de los autores de la serie GL-Conference, <http://www.textrelease.com/glpublications/glproceedings.html>. Otros autores han utilizado los parámetros cuantitativos para evaluar el impacto de LG en diferentes disciplinas.

Los resultados del trabajo de Di Cesare y Ruggieri (2006) confirman que la literatura gris es una importante fuente de información, incluso aunque se haya observado variabilidad de su uso e intensidad entre los distintos campos del conocimiento. Esta investigación también confirma que las herramientas bibliométricas utilizadas para el análisis de la literatura convencional tienen aplicaciones válidas en el análisis de la literatura gris. Sin embargo, los parámetros aún no se han codificado para la evaluación de impacto de LG en la literatura convencional en los

diferentes sectores científicos. Por último, el uso de este parámetro permite resaltar eventuales variaciones en el impacto con el tiempo. Este parámetro podría ser válido para la evaluación de impacto en sectores individuales y para el estudio de la evolución de impacto como resultado de los cambios en los procesos de difusión de la información. En consecuencia, se puede utilizar de manera rentable durante esta fase actual cuando el uso de Internet está abriendo progresivamente hasta los límites entre la literatura gris y convencional.

I.5.4. Las tesis doctorales. Las tesis electrónicas en acceso abierto

“Las tesis y disertaciones son los dos tipos más útiles de la producción científica invisible y la más invisible de los tipos de producción útil” (Suber 2012).

Las tesis doctorales se han incluido hasta ahora en la categoría de literatura gris, el término más difundido de todos los utilizados para denominar un conjunto de documentos de variada tipología que tienen como principal nota común la de no circular mediante los canales habituales de producción y comercio del libro, y que, por tanto, presenta problemas de identificación y accesibilidad (Orera Orera, 2003).

Son muchos los autores que han definido la tesis desde distintas perspectivas a lo largo del tiempo. Las tesis electrónicas también han sido definidas de varias formas por organizaciones nacionales e internacionales, universidades, instituciones, etc. explicando sus características, herramientas, beneficios y uso.

Orera Orera (2003) define la tesis doctoral como un documento que contiene los resultados de un trabajo de investigación que se presenta generalmente en una universidad, con el fin de obtener el grado de doctor. Todo ello, independientemente del diferente significado y finalidad que el doctorado tiene en los diferentes países, dependiendo del modelo de sistema universitario. La autora alude al creciente interés de las tesis debido a una serie de causas tanto

cuantitativas como cualitativas. En primer lugar hay que citar como una de estas causas, su importancia para el desarrollo de la investigación científica, por tratarse de trabajos de investigación que partiendo de informaciones anteriores, realizan aportes originales y abren nuevas líneas de trabajo para futuras investigaciones, lo que permite avanzar en el desarrollo del conocimiento. De forma genérica, entre las aportaciones originales hay que señalar el denominado repertorio bibliográfico final, que suele ser altamente selectivo, y recoge los trabajos que han servido de apoyo al investigador. Otra aportación importante lo constituye la metodología aplicada para el desarrollo de la investigación, que puede resultar de gran interés para investigadores de materias afines (Orera Orera, 2003).

La tesis doctoral es un trabajo de investigación riguroso, el trabajo por excelencia. Tiene que ser un trabajo inédito, de creación y por tanto, requiere que previamente se analice y estudie todo lo que se haya escrito, investigado y trabajado sobre el tema escogido (Arguimbau Vivó & Fuentes Pujol, 2009).

Las tesis doctorales constituyen un componente esencial de los estudios de postgrado y una importante fuente de publicación científica en las universidades (Echeverria, Stuart, & Blanke, 2015).

ISO (1986) define la tesis como el documento que presenta la investigación y las conclusiones de un autor, presentado por él en apoyo de su candidatura para obtener un grado o título profesional.

Según Orera Orera (2003) lo que hoy es un documento de investigación fue a veces en su origen un mero discurso protocolario. El desarrollo de la metodología científica y de normas para la elaboración de tesis las ha convertido en documentos de tipo científico, resultando cada vez más rigurosas e interesantes para el investigador. Desde el principio hasta ahora, las tesis doctorales

han experimentado una importante evolución, que se pone de manifiesto al estudiar la historia de las mismas en algunas universidades.

Paillassard, Schöpfel, & Stock, (2007) la definen como el documento de examen necesario para obtener el grado de doctor, y una pieza importante para la transferencia de conocimiento científico, económico y desarrollo cultural en el marco de trabajo del Área Europea de Educación Superior.

Las tesis son los trabajos de investigación por excelencia, son instrumentos para alcanzar un nuevo grado académico, son indicativos de la producción científica y son instrumentos de colaboración entre las universidades y las empresas (Arguimbau Vivó & Fuentes Pujol, 2009).

Las tesis doctorales, desde la perspectiva de la investigación, son excelentes documentos para caracterizar la investigación original en un determinado campo. La tesis doctoral es el primer trabajo científico, aquel que abre las puertas al mundo de la investigación, y seguramente, uno de los que más esfuerzos y recursos (sobre todo humanos y de tiempo) consume en la vida de un investigador. Pero, ante todo es, o debería ser, un trabajo de investigación original, esto es, debe aportar nuevo conocimiento. Y para acreditarlo se establece un sistema de control y de evaluación mediante un tribunal de personas, sabias en la materia, que lo certifica (Delgado López-Cózar, Torres-Salinas, Jiménez-Contreras, & Ruiz-Pérez, 2006).

La tesis doctoral es un trabajo de investigación realizado para obtener el Grado de Doctor, que lleva a cabo el doctorando en un departamento universitario o centro de investigación, bajo la orientación de un director de tesis (Miguel Alonso, 2000).

Para López-Yepes, Fernández Bajón y Prat Sedeño (2005) la tesis doctoral es el primer trabajo de investigación en el sentido más estricto, “por lo que debería ser el primer trabajo publicado en el tiempo” (esta idea ya es obsoleta, como se puede comprobar los programas de doctorado actuales). Para estos autores, la aprobación de la tesis supone la suficiencia como investigador.

Su objeto es suficientemente amplio de acuerdo con las características de este tipo de trabajo. Permite obtener el Título de Doctor.

Las motivaciones que conducen a una persona a realizar el doctorado son muy distintas, puesto que no será la misma motivación la de alguien que aspira alcanzar una plaza en docencia que la de otra persona cuyo interés es alcanzar un mayor grado académico o la de un tercero que pretende profundizar en determinada área del conocimiento o perfeccionar su desempeño en la investigación (Agudelo et al., 2003).

Todos los autores coinciden en afirmar que las tesis doctorales son un elemento fundamental en el desarrollo de la investigación y suponen un hito en la carrera académica de quienes realizan este importante trabajo. Por lo general, este tipo de estudios supone un avance en el progreso científico, por lo que la difusión de los mismos es de gran interés para docentes, profesionales e investigadores (Sorli Rojo & Merlo-Vega, 2002).

El acceso a las tesis doctorales no siempre ha resultado fácil, ya que no ha contado con una normativa de carácter general y no se ha regido por criterios uniformes (Moralejo Sánchez, 2000). Hasta la irrupción de la digitalización y de internet, las tesis doctorales solo se podían consultar a partir de ediciones impresas o microfilmadas, lo cual implicaba toda una serie de obstáculos para su diseminación y para que los resultados obtenidos tuviesen rápida incidencia en los ámbitos científicos. Todos estos problemas se han resuelto gracias a la variedad de formatos que permiten digitalizar un documento y a los distintos sistemas de acceso telemático mediante los que se pueden conseguir dichos documentos.

Aparte de su valor como investigación original, las tesis doctorales han sido consideradas tradicionalmente como literatura gris debido a su inaccesibilidad física (Jones & Andrew, 2005; Orera Orera, 2003).

A principios de los años 2000, Sorli Rojo y Merlo-Vega (2002) comentaban en su artículo sobre recursos en Internet de tesis doctorales que ya se había empezado a actuar contra el oscurantismo en el que habitualmente se habían visto rodeadas las tesis doctorales, como trabajos que apenas trascendían, a no ser que se editasen en forma de monografías, artículos o contribuciones a reuniones científicas, es decir, como trabajos derivados. Los autores afirmaban que en aquellos momentos la tendencia era poner a disposición de cualquier usuario de Internet los recursos que permitieran conocer la existencia de las tesis doctorales, e incluso el acceso al texto completo de las mismas. Como bien indican Sorli Rojo y Merlo-Vega (2002), la existencia de redes internacionales, como la *Networked Digital Library of Theses and Dissertations*, y proyectos nacionales, como el que coordina la Biblioteca Virtual Miguel de Cervantes, <http://www.cervantesvirtual.com/tesis/>, o el que realizan las universidades catalanas, <http://www.tesisenred.net/>, demostraba que se estaba avanzando en esta dirección, facilitando al investigador los textos íntegros de centenares de tesis de todo el mundo.

Las tesis electrónicas deberían estar accesibles libremente para servir mejor al progreso científico e informar a toda la comunidad científica sobre los resultados de la investigación. Técnicamente esto está apoyado en las universidades por sus repositorios institucionales y por algunos editores (Schöpfel, Prost, et al., 2014a).

El depósito de las tesis electrónicas (e-tesis) en repositorios permite que la investigación original realizada por los científicos principiantes sea visible, accesible y pueda ser utilizada por el conjunto de la comunidad investigadora, además de aumentar las estadísticas del repositorio. Las estadísticas indican que las tesis que están accesibles por esta vía son más usadas y citadas en la investigación actual (Troman, Jacobs, & Copeland, 2007).

Junto con los artículos de revistas, las tesis doctorales son el contenido más importante de los archivos abiertos. Más de la mitad de todos los repositorios institucionales del directorio

OpenDOAR contienen tesis y disertaciones. En julio de 2016, el motor de búsqueda científico BASE ofrecía más de 4,7 millones de tesis y disertaciones electrónicas (ETD) a través del protocolo *OAI-PMH*.; el portal DART-Europe daba acceso a cerca de 700.000 tesis de investigación de acceso abierto de 595 universidades de 28 países europeos; en España, desde TDR ya son consultables más de 80.000 tesis.

Según el estudio de Schöpfel, Zendulkova y Fatemi (2014) las tesis electrónicas cubren todas las disciplinas científicas y representan más del 10% de los resultados científicos nacionales. Como los artículos de revistas, informes, ponencias o documentos de trabajo, las tesis son resultados del trabajo de la investigación científica y son relevantes para los sistemas de gestión de la información científica (CRIS). Contienen información valiosa sobre proyectos de investigación e instituciones y aportan experiencias útiles para la gestión de conocimientos y habilidades de autores y organizaciones.

Las tesis doctorales electrónicas representan una parte muy significativa de las publicaciones académicas. Especialmente, las tesis doctorales contienen los resultados de al menos tres años del trabajo de investigación científica llevado a cabo en un laboratorio, grupo o instituto de investigación, facultad o empresa. A pesar de ello, las tesis doctorales son producidas por las universidades como una parte de la literatura gris científica, y consideradas como material de biblioteca, difundiéndose en un número limitado, con un estatus legal específico (Juznic, 2010; Schöpfel, Zendulkova, et al., 2014).

Suber (2015) defiende que las tesis y disertaciones son los dos tipos más útiles de la producción científica invisible y la más invisible de los tipos de producción útil. Debido a su alta calidad y baja visibilidad, valdría la pena resolver el problema para su acceso. Afortunadamente el acceso abierto a las tesis y disertaciones electrónicas (ETDs) es más fácil que para cualquier otro tipo de literatura científica. Esta facilidad viene dada principalmente por dos hechos: el primero que los

autores aún no han transferido ningún derecho de explotación a la editorial, por tanto no se ralentizan los procesos ya que no necesitan es necesario tramitar permisos ni negociar con los editores; y en segundo lugar, prácticamente todas las tesis en la actualidad nacen en formato digital, y las instituciones en las que se presentan en formato electrónico generalmente las distribuyen en acceso abierto, justo lo contrario que hacen las editoriales de revistas por defecto.

A lo largo de los años el porcentaje de tesis electrónicas que están en acceso abierto ha ido creciendo mundialmente en todas las disciplinas científicas.

Como se viene reclamando internacionalmente por los promotores y defensores del movimiento *Open Access* lo abierto debería ser por defecto y no una excepción. Sin embargo la realidad es diferente.

Incluso en los repositorios institucionales creados para proporcionar acceso a la producción científica de una institución académica y como un sector central de la denominada vía verde para el acceso abierto, falta las partes más o menos importantes de la producción científica. Esto es debido a la falta de concienciación, al embargo, al depósito de metadatos sin texto completo, al la confidencialidad del contenido, etc. Las tesis doctorales, en particular, se difunden con diferentes tipos de estado - algunas son de libre disposición, otras están bajo embargo, confidencialidad, acceso restringido a la comunidad (facultad, universidad) o no están disponibles de ninguna forma.

Al contrario de la filosofía del acceso abierto, muchos RIs no son totalmente Open Access (Prost & Schöpfel, 2014). Pero mientras que otros documentos pueden estar disponibles a través de canales alternativos (revistas, monografías, etc.), las tesis electrónicas no lo están la mayoría de las veces.

I.5.5. Beneficios y riesgos de las tesis en acceso abierto: percepción de los autores

En cuanto a la percepción de los doctorandos respecto a la difusión de las tesis en acceso abierto, Stanton y Liew (2011) comentan en uno de sus artículos que hay pocos estudios centrados específicamente en el papel del trabajo de los estudiantes en los repositorios. Existen dos artículos de Pickton y McKnight (2007; 2006) sobre las percepciones y necesidades de los estudiantes de postgrado con respecto al acceso abierto y a los repositorios institucionales. Los autores refieren que la mayoría de estudiantes están entusiasmados de difundir sus tesis y otros trabajos de investigación a través de los repositorios. El apoyo al principio del acceso abierto es el principal factor de motivación. Sin embargo, se encontró que los estudiantes al igual que el personal académico tenían las mismas preocupaciones sobre el acceso abierto y sobre el archivo en repositorios, particularmente en lo que respecta al plagio, la calidad y a la posible dificultad de su posterior publicación en revistas (Pickton & McKnight, 2006). En el segundo artículo los autores consideran la inclusión del trabajo de los estudiantes desde la perspectiva de los administradores de repositorios. A través de las respuestas de treinta y cinco instituciones, se extrae como conclusión que los gestores de repositorios creen mayoritariamente que el repositorio es el lugar idóneo para depositar la investigación de los estudiantes (Pickton & McKnight, 2007) aunque muchos consideraron que los trabajos que se depositaran en los repositorios deberían cumplir con ciertos criterios para garantizar el control de calidad. Algunos administradores de repositorios indicaron que estaban tratando de hacer frente a la falta de conciencia de los repositorios mediante la inclusión de información sobre el repositorio en la formación de las habilidades de investigación de postgrado.

El principal obstáculo parece ser el temor a que el hecho de difundir una tesis en acceso abierto pueda reducir la posibilidad de que una revista publique artículos derivados de la misma

(Ramírez et al., 2014; Ramírez, Dalton, McMillan, Read, & Seamans, 2013). Mientras que estos temores pocas veces podrían estar justificados, la evidencia demuestra que en la mayoría de los casos no lo están (Suber, 2012b, 2015), y que la disponibilidad de las tesis doctorales en Internet no impide que la gran mayoría de editoriales publique artículos derivados de la actividad investigadora contenida en dichas tesis (McMillan, 2001).

En los últimos años ha surgido un problema con las “revistas predatoras” derivado de la presión por “publicar o perecer” a la que se ven sometidos los jóvenes investigadores. Los estudiantes de doctorado y jóvenes investigadores, sobre todo en los países en vías de desarrollo, parecen ser la mayor parte de las víctimas de las revistas predatoras (Shaw, 2013; Xia et al., 2015). Estos investigadores están ansiosos por difundir sus publicaciones y ser reconocidos por la comunidad académica, por lo que son fácilmente atraídos por el modelo de publicación rápida y fácil que ofrecen estas revistas de acceso abierto. Presumiblemente, muchos autores que publican en estas dudosas revistas electrónicas son jóvenes investigadores del Hemisferio Sur y en particular del mundo musulmán (Truth, 2012; Xia et al., 2015).

Las universidades esperan que las tesis en abierto sirvan de ejemplo para las próximas generaciones de investigadores, demostrando lo fácil que resulta proporcionar el acceso abierto, los beneficios que conlleva, y cómo puede convertirse sin ningún esfuerzo en una rutina. Ayuda a crear hábitos de autoarchivo al ofrecer a los autores una audiencia potencial más allá del tribunal de tesis, una política de acceso abierto fortalece los incentivos existentes para hacer un trabajo riguroso y original. Si una universidad requiere que las tesis doctorales sean obras originales y significativas, entonces cabría esperar que se hicieran públicas, al igual que se espera que el resto de la producción científica se haga pública. Compartir las tesis y las disertaciones que cumplen estrictas normas académicas se refleja positivamente en la institución y en los beneficios que pueden obtener los investigadores en una disciplina. La misión de la universidad para que los jóvenes investigadores avancen en la investigación tiene dos facetas.

La primera consiste en ayudar a los estudiantes a producir un buen trabajo, y la segunda en ayudar a los demás a encontrar, usar y seguir construyendo sobre el buen trabajo realizado (Suber, 2012b, 2015).

Es destacable un caso de estudio presentado por Rodríguez-Márquez, Daoutis, y Gravelly (2014) sobre las necesidades de capacitación de estudiantes de doctorado relacionadas con la presentación de tesis electrónicas en la Universidad de Surrey de Reino Unido. Se consultó a los estudiantes de investigación, supervisores, administradores y altos directivos para evaluar el conocimiento de los temas relacionados con las e-tesis e identificar las áreas que debían ser abordadas en el programa de formación. Los autores hicieron hincapié en cuatro áreas principales en las que se necesitaba una mayor sensibilización y formación: la preparación de una tesis electrónica, y el conocimiento de cómo este formato crea nuevas oportunidades para la presentación de la investigación; el conocimiento del acceso abierto a las tesis y beneficios relacionados; la adopción de prácticas que se ocupan de los derechos de autor; y la formación en el uso de la plataforma de presentación electrónica. En 2015 en el Reino Unido la mayoría de las instituciones habían pasado a depositar sus tesis electrónicas. El principal impulsor de esta transición fue EThoS de la British Library, <http://ethos.bl.uk/>, a través de su servicio de digitalización y *kit* de herramientas de soporte asociado; por otro lado el informe Finch, fundamental en la formación de acceso abierto en el Reino Unido en los últimos años, también aboga por los repositorios institucionales para mejorar el acceso a las tesis.

Se considera que en las universidades españolas sería recomendable realizar este tipo de campañas de formación relacionadas con las tesis electrónicas en acceso abierto, sobre todo con el fin de informar y formar sobre las oportunidades que ofrecen los repositorios institucionales y la normativa vigente en relación con el depósito de las tesis doctorales en repositorios de acceso abierto, como por ejemplo el Real Decreto 99/2011 (Ministerio de Educación, 2011).

I.5.6. El aumento de las e-tesis en acceso abierto. El valor de las e-tesis en los repositorios institucionales

La digitalización de las tesis y el movimiento hacia el aumento de la accesibilidad a esta masa considerable de resultados de la investigación comenzó en la Universidad *Virginia Tech* en los EE.UU., dando lugar a la NDLTD en 1997, que ahora cuenta con miembros en todo el mundo (Vijayakumar & Vijayakumar, 2007). Estas iniciativas se han generalizado en los países desarrollados como Estados Unidos, Alemania, Canadá, Australia y el Reino Unido, pero los países en desarrollo también han adoptado los beneficios del acceso abierto, incluyendo China, Brasil, India y los países de África, a través de la Asociación de Universidades Africanas.

Como era de esperar, las motivaciones para tales desarrollos difieren. La filosofía de acceso abierto se basa en gran parte en el argumento de que fomentará las mejoras en los estándares académicos y maximizará el impacto de los descubrimientos científicos. Sin embargo, otra motivación destacada es que en el entorno competitivo en el que operan las universidades actualmente, el aumento de la visibilidad de las publicaciones y de los resultados de la investigación podría proporcionar una gran ventaja para obtener fondos destinados a la investigación así como atraer a estudiantes de postgrado, tanto nacionales como internacionales (JISC, 2008).

Entre las economías emergentes, sin embargo, los repositorios de tesis están siendo considerados como un medio importante para superar los obstáculos financieros que a menudo impiden a los investigadores el acceso a la información de la investigación que necesitan (Ghosh, 2009).

Por ejemplo, es un hecho reconocido que en las universidades africanas los investigadores tengan grandes dificultades para acceder al trabajo de otros investigadores, y que por tanto el

desarrollo de repositorios en estos países podría hacer posible la investigación para lograr colectivamente una masa crítica y así mejorar la visibilidad de los investigadores a nivel internacional (Moyle, Barnes, Brown, & Sadler, 2012).

Las tesis electrónicas representan un segmento creciente del acceso abierto, cuyo contenido está disponible en repositorios institucionales donde contribuyen al impacto y al posicionamiento de la institución en los *ranking*. Más de la mitad de todos los repositorios institucionales listados en el directorio OpenDOAR contienen tesis electrónicas. La mayoría de estos proyectos de acceso abierto presentan funcionalidades comunes, como el acceso al texto completo y el cumplimiento de OAI-PMH. Pero también presentan diferencias con respecto a los metadatos, a las políticas, a las restricciones de acceso, a la representatividad, a los formatos de los ficheros, al estatus, a la calidad y a los servicios relacionados con los repositorios.

Como se ha explicado anteriormente, una de las definiciones de los repositorios institucionales es que son herramientas para recoger, almacenar y diseminar los resultados académicos dentro y fuera de la institución (Jain, 2011), es decir son un conjunto de servicios para la gestión y diseminación de materiales digitales creados por la institución y los miembros de su comunidad (Lynch, 2003), son archivos abiertos que sirven a los intereses de la universidad – investigadores y profesores- mediante la reunión de sus resultados intelectuales para su acceso a largo plazo, preservación y gestión (Swan & Carr, 2008). Una de sus principales características es la de su gran diversidad. Los repositorios institucionales tienen diferentes políticas, procedimientos, funcionalidades, servicios y metadatos, y contemplan diferentes modelos de negocio y estrategias para su financiación.

Como se ha venido comprobando en los repositorios institucionales el corpus de tesis electrónicas es cada vez más importante. En la Figura 23 se muestra el gráfico de evolución de las tesis en RIs, desde 2001 hasta 2016.

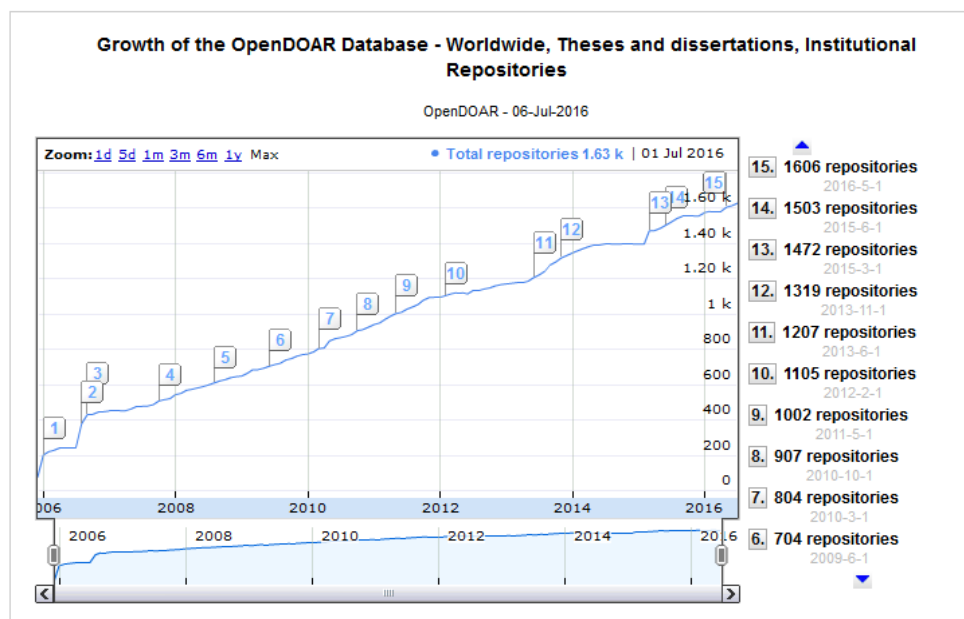


Figura 23. Crecimiento mundial de tesis electrónicas en los repositorios institucionales. Fuente: OpenDOAR

En julio de 2016, el directorio internacional OpenDOAR listaba más de 1.600 RIs con tesis electrónicas, representando más de la mitad de los repositorios registrados en este directorio. A veces algunas instituciones han visto en los RIs una forma para preservar y difundir sus tesis doctorales. Después de los artículos preprint o postprint, las tesis electrónicas son el tipo de documento más importante de los archivos abiertos, significativamente más importante, por ejemplo, que los documentos de trabajo, informes, capítulos de libros o *proceedings* de conferencias.

Es difícil estimar el número exacto de tesis electrónicas que están disponibles actualmente en repositorios institucionales. El catálogo colectivo de la NDLTD contiene actualmente, en 2016, más de cuatro millones de e-tesis, cuya mayoría probablemente estén depositadas en repositorios institucionales. Esto demuestra claramente el éxito de la ruta verde del acceso abierto para la información científica y, en nuestro caso, para la literatura gris científica.

Sassi et al. (2010) declararon que el acceso abierto es la llave al conocimiento y que las facilidades para compartir de la web y el acceso distribuido a la información de manera abierta

son herramientas clave para la literatura gris. Para los autores los repositorios institucionales se convertirán en el futuro de la literatura gris, y la literatura gris apenas existirá más allá del acceso abierto.

Sin embargo “abierto” no es suficiente. Para incrementar el acceso y el impacto no es suficiente con subir el contenido a los servidores. Por lo menos es necesario un mínimo de valor añadido (Schöpfel et al., 2011). Las cuestiones irán dirigidas entonces a cómo pueden las instituciones dar valor a las e-tesis y cómo pueden incrementar la calidad del contenido y la provisión de servicio en un entorno de acceso abierto.

Schöpfel (2013) señala cinco vías para añadir valor al depósito y disseminación de las tesis electrónicas y describe algunos desarrollos que estaban cambiando los repositorios institucionales. La revisión llevada a cabo por el autor ofrece muy diferentes soluciones: la Red Peruana de Tesis Digitales, con un único punto de acceso para todos los tipos de tesis, como el servicio EThOS de la British Library, los programas de tesis electrónicas locales, como el del campus Emory de la Universidad de Atlanta, los grandes servidores de documentos, tales como e-doc en la Universidad Humboldt de Berlín, y las soluciones internacionales para repositorios, como el portal europeo de tesis doctorales electrónicas DART. Todas estas propuestas tienen en común dos funcionalidades: una que se trata de repositorios institucionales de acceso abierto, y otra que contienen tesis electrónicas. Sin embargo, cuando se comparan entre sí, se pueden identificar cinco características específicas que constituyen una diferencia notable, y proporcionan la excelencia científica.

Estas características son: calidad, metadatos, formato, interoperabilidad y servicios.

En cuanto a la calidad, según Macha y Jager (2011) el contenido es el factor más importante que ha sido citado por los investigadores para mostrar el éxito de un repositorio. Un buen repositorio institucional no solo se define por un conjunto de estándares y criterios para la selección y

validación de los depósitos sino que también comunica y promueve esta política editorial. Para las tesis electrónicas esto significa únicamente versiones validadas con revisiones menores si son necesarias por el tribunal de la tesis, pero sin preprints o autoarchivos no controlados. Esto también significa que el repositorio institucional forma parte del flujo de trabajo local de las tesis electrónicas o que incluye un punto de la validación formal e institucional durante el proceso de depósito. Independiente de la validación institucional, adicionalmente las normas del repositorio institucional podrían contener algunos requisitos mínimos para la aceptación, como el acceso sin restricciones a texto completo, la cesión de los derechos, o limitar la aceptación únicamente a las tesis de investigación. Un sistema de revisión específico podría ser útil para asegurar la calidad. Los mismos criterios de calidad selectiva deberían aplicarse a la digitalización retrospectiva de las tesis impresas cuando el RI contenga materiales tanto nacidos digitales como materiales más antiguos reutilizados.

Una segunda característica que marca la diferencia es la descripción del contenido y contexto de los ficheros de las e-tesis, es decir los metadatos. En la práctica existen variaciones considerables. Muchas de estas variaciones implican cualificadores locales y el uso de elementos de metadatos para registrar la información que no se corresponde con el significado establecido de los elementos (García-Peñalvo, Merlo-Vega, et al., 2010; Park & Richard, 2011). Una buena y rica descripción bibliográfica aumenta la capacidad de búsqueda de los documentos depositados y permite el análisis bibliométrico, la edición de libros (impresión bajo demanda) y otras funcionalidades y servicios.

En general los metadatos de tesis electrónicas en repositorios institucionales deberían ser ricos y estandarizados, y el uso de estándares debería ser promovido por buenas prácticas. Los estándares de metadatos son especialmente importantes para instituciones de tamaño pequeño o medio y en redes como DART. En este sentido DART utiliza la versión simplificada de *Dublin Core* como formato básico, mientras que la infraestructura nacional francesa de tesis

electrónicas utiliza el formato más sofisticado TEF (Ducloy et al., 2006). Otros estándares de metadatos utilizados son los de la NDLTD para DC/MARC, el esquema UMI XML DTD de ProQuest (C. Marsh & McLean, 2008) y el EVSKP-MS en la República Checa. Sea cual sea el esquema de metadatos que se implemente, deberá ser flexible, facilitar la conversión a otros formatos y deberá cambiar con las nuevas políticas (Heyse, 2007).

En cuanto al formato, los repositorios institucionales ofrecen diferentes formatos de archivo según el uso y la finalidad. Los RIs contienen texto completo, no solo metadatos. Los formatos de depósito tienen que ser susceptibles de búsqueda, abiertos, y apropiados para la preservación a largo plazo y la explotación del contenido (revisión, bibliografía, etc.). Las tesis de ciencias tendrían que ser vistas no como simples objetos destinados a la preservación y al archivo, sino como las únicas fuentes que contienen potencialmente datos de valor que podrían ser extraídos y reutilizables (Morgan et al., 2008). Los formatos estándar para el depósito y la disseminación son MS Office Word y Adobe PDF. Pero algunos repositorios aceptan y procesan diferentes formatos como XML, LaTeX, Postscript o RDF, más recomendable para algunas disciplinas, la consulta semántica y la reutilización. Para la digitalización retrospectiva de las tesis, es necesario OCR y la producción de archivos de texto antes de su presentación.

La cuarta funcionalidad es la interoperabilidad. Para incrementar el descubrimiento y la disponibilidad de las e-tesis, los repositorios deberían estar en red e interconectados. Esta interconexión puede darse a diferentes niveles, por ejemplo regional, nacional o internacional. *La Networked Digital Library of Theses and Dissertations* es el mejor ejemplo de red internacional.

Como redes nacionales serían destacables *TDR* en España, *ETHOS* en el Reino Unido -red nacional alojada por la British Library- y el sistema *STAR* en Francia. Estas y otras iniciativas presentan las principales características y condiciones para el trabajo en grupo y la interconexión: un formato estándar compartido para el archivo, el intercambio y la recolección

(OAI-PMH), una política explícita, como el compromiso con la red y la voluntad colectiva para el intercambio y la interconexión, una amplia arquitectura de estado (regional, nacional, etc.), y un único centro para la investigación publicada en las tesis doctorales.

Y por último la funcionalidad servicios. Sin duda, cualquier repositorio institucional ofrece algunas funcionalidades básicas, tales como los servicios de búsqueda simple y avanzada, o las diferentes opciones de navegación, visualización y descarga. Sin embargo si la institución que aloja el servicio o el proveedor de servicios quiere incrementar el uso y el impacto de su producción y contenido, deberían desarrollar servicios y prácticas innovadores y periféricos de acuerdo a lo que Halbert (2008) llama un proceso centrado en el usuario de priorización en el desarrollo de servicio. En un estudio sobre repositorios abiertos de Bester (2010) se distinguen nueve opciones de búsqueda y diez opciones para la personalización y gestión de la referencia, como la exportación de listados bibliográficos en diferentes formatos, incluido XML. Otros desarrollos interesantes en este sentido son:

- Herramientas de la web social: interactividad, foros de discusión y comentarios (Cocciolo, 2010; Millard et al., 2010; Waddington et al., 2012).
- Búsqueda federada y herramientas de descubrimiento más sofisticadas (Dowling & Steans, 2011).
- Uso de estadísticas, métricas y citas (Huaroto, 2008; Walker, 2011).
- Impresión bajo demanda en formato libro (Rajendiran, Ramesh Babu, & Gopalakrishnan, 2005).
- Opciones para la protección de derechos de autor o licencias *Creative Commons* (Hagen, 2007).
- Preservación en múltiples copias (Mikeal, 2008).
- Mapeo para transferencia de metadatos (S. Deng & Reese, 2009).

Algunos de estos desarrollos son posibles incluso con un modesto equipo de programación disponible para desarrollar proyectos *ad hoc* (Halbert, 2008). En un nivel más general, los servicios de los RIs para la difusión de las e-tesis deben ser flexibles, con una capacidad de adaptación rápida; el *software* debe ser fácil de usar y fiable, pudiendo estar abierto a otros proveedores de servicios y / o integrado en otro entorno de servicio, tales como el e-learning (García-Peñalvo, 2008d; García-Peñalvo & Seoane-Pardo, 2015).

Estas cinco funcionalidades descritas añaden valor a la difusión de las tesis doctorales vía repositorios institucionales. Todas estas herramientas ya han sido probadas e implementadas en muchos casos pero no están generalizadas, dependiendo sobre todo de las decisiones políticas y técnicas de las instituciones en relación al tratamiento de las e-tesis depositadas en repositorios institucionales de acceso abierto.

Además de estas funciones, se distinguen otros desafíos -comunes a toda publicación científica- para las e-tesis depositadas en repositorios institucionales: la orientación hacia el futuro y anticipación a la transformación de la comunicación científica.

En cuanto a los datos de las e-tesis, el formato de texto debe permitir la extracción y reutilización de contenido. De manera más general, las tesis doctorales se depositan regularmente acompañadas de material complementario, por ejemplo, multimedia, imágenes, hojas de cálculo, música, etc. En el entorno digital de los repositorios abiertos y de servicios de valor añadido, este material es en una rica fuente de resultados de investigación y de conjunto de datos. Pueden servir para mejorar tanto la investigación como el acceso. Un acceso profundo para la investigación podría incluir datos, representaciones visuales y auditivas y enlaces a material de apoyo (White, 2007).

Según Schöpfel (2013) el impacto de la diseminación de las tesis doctorales a través de repositorios institucionales es al menos triple:

1. El RI debería soportar otros formatos no textuales como avi para vídeos o mp3 para archivos de música.
2. El *host* del RI debería desarrollar un entorno de servicios innovadores, extendiendo los servicios básicos y de valor añadido para los archivos de texto a los conjuntos de datos (*datasets*).
3. La institución necesita revisar las condiciones legales del depósito y diseminación de los conjuntos de datos y de otros materiales. Algunos contenidos podrían estar protegidos por intereses privados o de confidencialidad. Por otra parte la institución y/o el autor deberían atribuir una licencia *Creative Commons*, o similar, que permita la máxima reutilización y explotación de los archivos de las tesis doctorales.

Si aplicamos la lógica de la e-ciencia, el material suplementario de las tesis doctorales no debería estar disponible solamente como un apéndice o anexo sino ser también extractable y reutilizable cuando no está unido a la tesis, como un *dataset* independiente e interconectado a otros datos (Ubogu & Sayed, 2008).

No podemos dejar de mencionar en esta investigación la conexión de los RIs a los *Current Research Information Systems* (CRIS). Actualmente una de las implementaciones en las que trabajan no solo las universidades a nivel internacional sino la mayoría de las universidades españolas es la conexión a de los RIs a los CRIS. Las tesis doctorales en repositorios institucionales deben contribuir a la evaluación de la producción científica en el entorno emergente de los CRIS.

Según el informe final de EUNIS y EUROCRIS en el que se analiza la situación de los CRIS en Europa, el 62% de las instituciones tienen los dos sistemas: CRIS y RI. Un 96% de los repositorios institucionales almacenan principalmente metadatos y el texto completo de las publicaciones y un 86% tesis y disertaciones. El *software* más utilizado para los repositorios

institucionales es DSpace (56%). La mayoría de las instituciones tienen vinculados el CRIS y el RI (Ribeiro, Castro, & Mennielli, 2016).

Una buena gestión del repositorio puede simplificar la tarea de los investigadores y sus respectivos departamentos, y servicios de administración. Las publicaciones son elementos esenciales de los sistemas CRIS. Los repositorios contienen metadatos de los resultados de la investigación de las instituciones y podrían estar disponibles para los CRIS con el fin de monitorizar y evaluar la producción científica y las líneas de investigación (Lambert et al., 2005; Ribeiro et al., 2016).

I.5.7. Diseminación de las tesis electrónicas. Redes, recolectores y portales de e-tesis

Las tesis junto con artículos de revistas son el contenido más importante de los repositorios. Los repositorios de e-tesis se han convertido en habituales en las universidades de todo el mundo. Los datos sugieren un importante incremento del uso y la citación de las tesis doctorales en la actividad de investigación. Más de la mitad de los repositorios institucionales de todo el mundo contienen tesis según los datos del directorio *OpenDOAR*.

En cuanto a los aspectos técnicos y políticos, se considera que la interoperabilidad de los repositorios y las políticas de mandato afectan positivamente al aumento de la visibilidad y de citación de los trabajos depositados en acceso abierto en los repositorios.

Los principales servicios de diseminación de las tesis doctorales, que son posibles gracias a la interoperabilidad de los repositorios, son los que a continuación se relacionan en este epígrafe. Se trata de redes, recolectores de repositorios y portales de e-tesis tanto nacionales como internacionales.

Networked Digital Library of Theses and Dissertations (NDLTD)

NDLTD es una organización internacional dedicada a promocionar la adopción, creación, uso, diseminación y preservación de las tesis electrónicas (e-tesis). Apoyan la publicación electrónica y el acceso abierto a la ciencia con el fin de mejorar el intercambio de conocimiento en todo el mundo. El portal de NDLTD incluye recursos destinados a administradores de universidades, bibliotecarios, profesores, estudiantes, y público en general. Los tópicos o temas incluyen: cómo encontrar, crear y preservar e-tesis; cómo establecer un programa de e-tesis; aspectos técnicos y legales; y las últimas novedades e investigación en la comunidad de las e-tesis.

Las metas de esta organización son principalmente tres:

- Ser la principal organización internacional para la promoción de las e-tesis en todo el mundo.
- Proporcionar recursos útiles e innovadores, estándares y tecnología para el desarrollo de programas de e-tesis.
- Animar a las instituciones de educación superior para utilizar los recursos de NDLTD y participar en sus actividades.

Entre sus actividades se encuentran:

- El soporte a las instituciones de educación superior para desarrollar sus propios programas de e-tesis mediante la adopción del envío, la recolección, y el archivo de las tesis y disertaciones electrónicas a sus propios repositorios y bibliotecas digitales y repositorios internacionales.
- El apoyo al movimiento Open Access.
- La sensibilización a las instituciones de educación superior sobre los beneficios de las e-tesis, incluyendo: la facilitación del proceso de escritura para los estudiantes, el aumento de la velocidad de transferencia de métodos de investigación y los resultados, la mejora

de la distribución de los métodos de investigación y resultados a través de la publicación electrónica y archivo, y los costes reducidos de impresión, procesamiento y almacenamiento.

En cuanto a la historia de la organización hay que resaltar que el concepto de tesis y disertaciones electrónicas, en inglés ETDs, fue discutido por primera vez en 1987 con ocasión de la conferencia organizada por UMI en Ann Arbor, Michigan, a la que asistieron representantes de *Virginia Tech*, Universidad de Michigan, y dos pequeñas empresas de *software*, Toronto-based SoftQuad y Michigan-based ArborText.

El proyecto permaneció dormido durante unos años, hasta 1991, cuando algunos profesores de Virginia Tech comenzaron a colaborar en el proyecto ETD investigando los problemas con la producción, el archivo y el acceso a las e-tesis. A comienzos de los años 90, se llevaron a cabo una serie de reuniones sobre el diseño del proyecto, en estrecha colaboración con la Coalition for Networked Information (CNI), Council of Graduate Schools (CGS), UMI, y otros grupos interesados. Al mismo tiempo, el proyecto Virginia Tech University Library's Scholarly Communications desarrolló procedimientos y sistemas de procesamiento, archivo, y acceso público para las obras de investigación de los graduados de *Virginia Tech*.

En 1993, el interés por las e-tesis se extendió con el comienzo de Monticello Electronic Library Project (*MEL*), apoyado por las universidades de Southeastern Universities Research Association (SURA) y la Southeastern Library Network (SOLINET, actualmente LYRISIS).

El resultado de varios años de intensa colaboración fue el *software* ETD que surgió de Virginia Tech en 1996, proporcionando un paquete completo de e-tesis de las universidades involucradas en el proyecto además de otras universidades del sureste de los Estados Unidos, tales como Auburn, Clemson, Universidad de Delaware, Universidad de Georgia, Georgia Tech, Oklahoma

State, Mississippi State, North Carolina State, y Universidad de West Virginia, que ayudaron a probar el *software*.

Desde 1996 el *software* está disponible para las instituciones de todo el mundo. Manteniendo su papel de liderazgo, Virginia Tech también coordinó el desarrollo e implementación de un sistema de biblioteca digital distribuido, de manera que se puede acceder fácilmente a las e-tesis de todas las instituciones participantes se puede acceder fácilmente. El sistema desarrollado permitía la navegación y la búsqueda basada por la institución, fecha, autor, título, palabras clave y texto completo, así como la descarga para la lectura local o la impresión de las e-tesis de todo el mundo.

Este esfuerzo inicial para crear una biblioteca digital mundial de e-tesis proporcionó el marco conceptual para lo que se convirtió en 1996 en la Networked Digital Library of Theses and Dissertations.

En 1998, las instituciones interesadas comenzaron a reunirse anualmente, lo que dio lugar a una serie de simposios sobre tesis electrónicas patrocinados por NDLTD y diseñados para ayudar a las universidades a iniciar proyectos de e-tesis. El primer simposio se celebró en la Universidad de Tennessee Health Science Center, y desde entonces, las reuniones han tenido lugar en las universidades de América del Norte, Australia y Europa, donde el primer lugar de celebración fue la Universidad Humboldt de Berlín, en 2003.

Actualmente, NDLTD incluye como miembros a cientos de universidades de todo el mundo además de las organizaciones asociadas, tales como Adobe, American Library Association, Association of Research Libraries, Coalition for Networked Information, Joint Information Services Committee, OCLC Online Computer Library Center, Proquest/UMI, y Theses Canada. Todos

ellos trabajando con el objetivo de desbloquear los beneficios del conocimiento compartido para todos.

TDX – TDR

Tesis Doctorals en Xarxa (TDX) o Tesis Doctorales en Red (TDR) es un repositorio cooperativo que contiene, en formato digital, tesis doctorales leídas en las universidades de Cataluña y otras comunidades autónomas españolas. La consulta de las tesis es de acceso libre y permite realizar búsquedas en el texto completo, por autor/a, director/a, título, palabras clave, universidad y departamento donde se haya leído, año de defensa, etc. Los objetivos de este repositorio, gestionado y coordinado por el Consorci de Serveis Universitaris de Catalunya (CSUC), y patrocinado por la Generalitat de Catalunya, son:

- Difundir, por todo el mundo y a través de internet, los resultados de la investigación universitaria.
- Ofrecer a los autores de las tesis una herramienta que incrementa el acceso y la visibilidad de su trabajo.
- Favorecer la edición electrónica en abierto de la producción científica propia.

Son las propias universidades las que se encargan de difundir el repositorio entre sus estudiantes de doctorado para que, una vez presentadas y aprobadas sus tesis, entreguen la información necesaria para poderlas introducir en TDR. Desde el año 2007 se llevan a cabo también proyectos de digitalización para incorporar en el repositorio tesis de las que solo se conserva el original en papel o en microficha. A través de estos proyectos de digitalización ha sido posible reconvertir a formato digital más de 1.400 tesis. En 2008 y 2009 TDR recibió una ayuda de la Dirección General del Libro, Archivos y Bibliotecas del Ministerio de Cultura para llevar a cabo la digitalización de tesis.

Los derechos del autor/a de la tesis quedan protegidos a través de una declaración en la que se determinan los derechos que se reserva o cede. En la mayoría de los casos las tesis están sujetas a licencias *Creative Commons*. Las condiciones de uso se especifican en el registro de cada tesis (dentro de los metadatos).

TDR nació fruto del convenio *La Universitat Digital a Catalunya 1999-2003*, firmado el 8 de septiembre de 1999, por los entonces Comisionados para la Societat de la Informació y para Universitats i Recerca, la Universitat de Barcelona, la Universitat Autònoma de Barcelona, la Universitat Politècnica de Catalunya, la Universitat Pompeu Fabra, la Universitat de Girona, la Universitat de Lleida, la Universitat Rovira i Virgili y la Universitat Oberta de Catalunya, los entonces Fundació Catalana per a la Recerca, CBUC y CESCO. Posteriormente, se han incorporado la Universitat Jaume I (julio 2002), la Universitat de les Illes Balears (diciembre 2002), la Universitat de València (abril 2003), la Universitat Ramon Llull (octubre 2004), la Universidad de Cantabria (marzo 2005), la Universidad de Murcia (julio 2005), la Universitat de Vic - Universitat Central de Catalunya (mayo 2007), la Universitat Abat Oliba (diciembre 2007), la Universitat Internacional de Catalunya (enero 2008) y la Universitat d'Andorra (abril 2013).

La Universidad de Salamanca participa también en este portal a través del protocolo OAI-PMH, siendo recolectada la totalidad de la comunidad de Tesis Doctorales del Repositorio Gredos³¹, cuyo contenido actualmente asciende a más de 2.100 tesis.

Desde el año 2011 TDR participa en la cooperativa MetaArchive³² con el objetivo de asegurar la preservación de las tesis y de estar dentro de una de las iniciativas pioneras en el mundo en este ámbito. Esta preservación se lleva a cabo a través del *software* LOCKSS³³.

³¹ Tesis doctorales (Gredos), <http://gredos.usal.es/jsui/handle/10366/4756>

³² MetaArchive, <http://www.metaarchive.org/>

³³ Software LOCKSS, http://www.adpn.org/wiki/LOCKSS_Software

Las tesis incluidas en TDR están descritas con metadatos Dublin Core y siguen el protocolo de interoperabilidad *OAI-PMH* hecho que permite incrementar su visibilidad al ofrecerse conjuntamente con otros repositorios, como por ejemplo en *DART-Europe E-theses Portal* o en *OpenAIRE*³⁴. A través del uso de este protocolo TDR ofrece también, desde mayo de 2007, la opción de hacer una búsqueda conjunta de tesis españolas independientemente del repositorio en el que están incluidas.

DART- Europe

DART-Europe es una asociación de bibliotecas de investigación y de consorcios bibliotecarios que trabajan conjuntamente para la mejora del acceso global a las tesis doctorales europeas. DART-Europe está apoyado por LIBER³⁵ (Ligue des Bibliothèques Européennes de Recherche) y es el Grupo Europeo de Trabajo de la Networked Digital Library of Theses and Dissertations (NDLTD).

Los socios de DART-Europe pretenden facilitar a los investigadores un único portal europeo para el acceso a tesis electrónicas, y trabajan en la promoción de futuros desarrollos europeos sobre este tipo de documentos. Ofrece a sus socios un foro europeo para cuestiones relacionadas con las tesis electrónicas y puede facilitar la posibilidad de optar a financiación colaborativa destinada a hacer realidad la visión de DART-Europe entorno a las tesis electrónicas.

DART-Europe se nutre de las contribuciones de sus socios. Los socios apoyan los siguientes principios:

1. DART-Europe estimulará la creación, el descubrimiento y el uso de tesis electrónicas europeas, y mantendrá un portal centralizado para la agregación de tesis electrónicas y para su acceso.

³⁴ OpenAire, <https://www.openaire.eu/>

³⁵ LIBER, <http://libereurope.eu/>

2. Las bibliotecas y consorcios europeos están invitados a contribuir con metadatos al portal de DART-Europe. Los contribuyentes determinarán los términos y condiciones bajo las cuáles sus metadatos son incluidos.
3. DART-Europe agradece la contribución de socios para apoyar la gestión, descubrimiento, potenciar el uso y la preservación de tesis electrónicas, y para fomentar las metas y objetivos de DART-Europe.
4. Los socios designarán una persona de contacto con DART-Europe, y nombrarán como mínimo un representante para la participación en las listas de correo electrónico de DART-Europe.
5. DART-Europe agradecerá las ofertas de los socios para alojar las reuniones del consejo de DART-Europe, cuando la ocasión lo requiera.
6. Los socios ayudarán a consolidar la posición de DART-Europe como red internacional de excelencia en información, experiencia y recursos referentes a tesis doctorales electrónicas.
7. DART-Europe estará administrado por UCL (University College London) y dirigido por un consejo constituido por representantes de las organizaciones asociadas. La constitución y bases de funcionamiento del Consejo de gobierno serán determinadas y periódicamente revisadas por el consejo.

Open Access Theses And Dissertations (OATD)

OATD.org es un portal que recoge las tesis electrónicas en acceso abierto de todo el mundo.

Los metadatos sobre las e-tesis proceden de más de 1.100 facultades, universidades, e instituciones de investigación de todo el mundo. OATD indexa, en agosto de 2016, más de 3.251.000 tesis.

El repositorio Gredos también es recolectado por OATD vía OAI-PMH, participando con la totalidad del contenido de la comunidad de Tesis doctorales como hace con los anteriores portales de e-tesis aquí referenciados.

I.5.8. Acceso a las tesis electrónicas: mejoras para el acceso

En el siglo XXI la Sociedad es cada vez más consciente de la importancia de la información y de su integración en la vida cotidiana, estando la calidad de vida estrechamente relacionada con los datos y con la información.

Los académicos juegan un papel significativo en el desarrollo de una nueva generación de científicos e investigadores y de mejores situaciones para afrontar los cambios tanto a nivel nacional como internacional. Indudablemente el sistema de Educación Superior contribuye a que la mayor parte de este escenario sea posible y el ingrediente más importante podrían ser las tesis doctorales, que son los trabajos de investigación de los individuos sobre un tema en particular que contribuyen directa o indirectamente al crecimiento y desarrollo de la sociedad.

Algunos autores han trabajado sobre el tema de cómo mejorar el acceso a las tesis doctorales electrónicas (Deng & Reese, 2009; Farace, Frantzen, & Schöpfel, 2006; Fox, 2014; Schöpfel, Prost, et al., 2014a; Schöpfel & Soukouya, 2013; Shen, Goncalves, & Fox, 2013; Vijayakumar & Vijayakumar, 2007).

Para Vijayakumar y Vijayakumar (2007) desafortunadamente no existe un estándar uniforme a seguir donde la tecnología de la información y la comunicación predominen y allanen el camino para el acceso fácil y multiusuario con información exhaustiva a través de redes informáticas basadas en un modelo de acceso abierto. No obstante, existen esfuerzos esporádicos realizados en algunas partes del mundo como EE.UU., Reino Unido, Alemania, Australia y otros países avanzados para atender los problemas relacionados con la obtención del texto completo de las

tesis doctorales con el fin de llegar a su componente intelectual en todo el mundo ahorrando tiempo y evitando la duplicación de esfuerzos para enriquecer los proyectos multidisciplinares.

Para Fox (2014) lo esencial para mejorar el acceso será disponer de información adecuada para darle soporte. Hay dos tipos principales de información que se necesitarán mejorar para dar un buen acceso a las tesis: las tesis electrónicas en sí y los metadatos relacionados.

Es necesario también un mejor acceso a los procesos y servicios. En cuanto a satisfacción de los usuarios, sería deseable la búsqueda en el texto completo del contenido de las tesis electrónicas, para poder localizar tanto un capítulo como una obra completa. En este sentido, es fundamental la búsqueda facetada y la navegación, donde cualquier metadato pueda ser información susceptible de ser buscado y recuperado.

Por otra parte, se podrían ofrecer otros servicios como la recomendación, la recuperación de imagen, de cita, y el soporte de hipertexto que permitiera una mejora en la navegación. Los investigadores deberían poder encontrar todas las e-tesis de un director en concreto, o de un grupo de investigación en particular, o la e-tesis o artículo cuya cita fuera particularmente influyente. También se necesita el depósito y procesamiento en repositorios institucionales de todas las e-tesis. Con el aumento de las e-tesis que se están escaneando del papel o microforma, son primordiales el OCR y la corrección relacionada de este para permitir una extracción de metadatos exitosa.

Si lo que pretenden las tesis electrónicas es mejorar la educación de postgrado y ampliar la colaboración de la investigación mundial, las metas a alcanzar según Fox (2014) deben ser:

- Ayudar a los estudiantes a comunicar con más efectividad.
- Conseguir que todas las tesis sean electrónicas.

- Colaborar para hacer que todas las tesis electrónicas estén en acceso abierto, accesibles y adecuadamente preservadas.

El núcleo del esfuerzo en la creación de tesis electrónicas requiere el uso de sistemas adecuados de autoría, herramientas y métodos. Afortunadamente mucho de este esfuerzo es posible hoy en día gracias al desarrollo de métodos de edición electrónica, de los que puede beneficiarse las tesis electrónicas. También es importante el desarrollo de técnicas para la creación y gestión de diversos tipos de contenido (por ejemplo los datos y la información) que forman parte de la investigación. Desde el mundo de las bibliotecas han llegado el *software* y los métodos para crear y mejorar los metadatos, con el fin de llevar a cabo una completa descripción de los productos de la investigación. Desde el mundo de las tesis electrónicas y de los repositorios han llegado diferentes sistemas que soportan la carga de las tesis con sus fases de envío, aprobación, y refinamiento integradas en adecuados flujos de trabajo. Normalmente estos se abordan a través de sistemas de gestión de contenidos, proporcionados por consorcios, instituciones nacionales o socios corporativos. A veces también se fomenta el intercambio, la difusión, la investigación y el descubrimiento de investigación de interés. A través de OAI-PMH, el repositorio local se convierte en proveedor de metadatos alimentando a las instituciones regionales, nacionales o internacionales; y a continuación estas proporcionan una serie de servicios, tales como el acceso y la preservación, así como otros productos de valor añadido. En este escenario las instituciones activas a menudo también digitalizan sus archivos antiguos, para ampliar la cobertura histórica más allá de de las obras recientes nacidas digitales. Existe una amplia variedad de servicios de bibliotecas o repositorios digitales que se pueden implementar como parte de un plan de acción de mejora de tesis electrónicas (Figura 24).

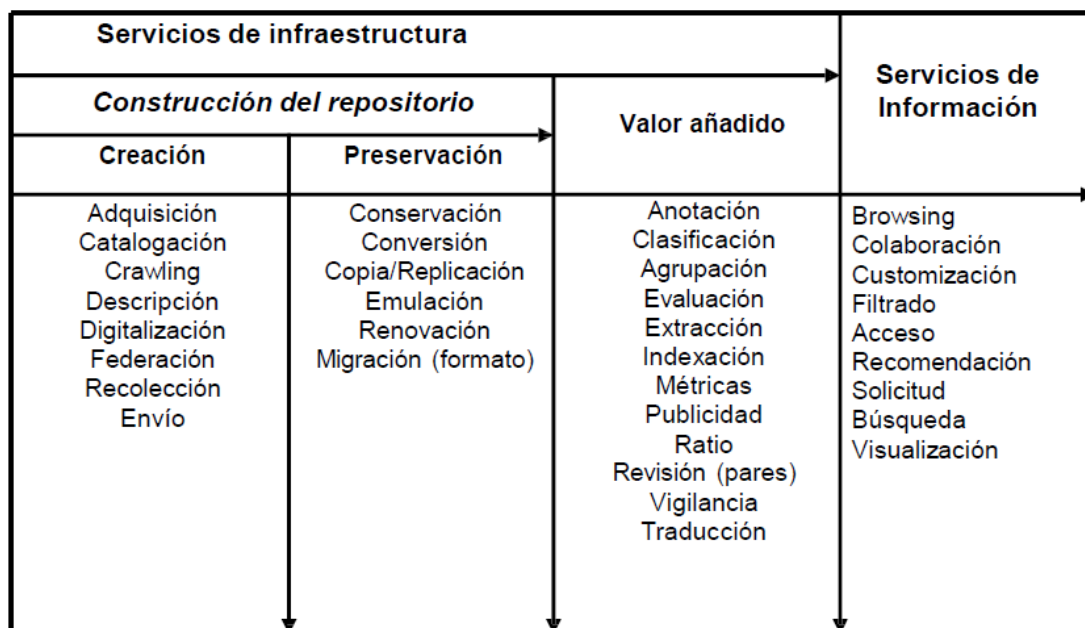


Figura 24. Servicios de infraestructura del repositorio. Fuente: Modificado de Fox, E. A. (2014). Improving the ETD landscape

Fox (2014) sostiene que la comunicación académica se basa en la calidad ya que el objetivo general es la mejora. En cuanto a la calidad y la mejora de las tesis electrónicas, Shen, Goncalves, y Fox (2013) exponen que la calidad de las tesis (junto con sus metadatos y otros objetos auxiliares) se puede evaluar en todas las fases de su ciclo de vida: creación, distribución, búsqueda y utilización de acuerdo a lo expresado en la Figura 24.

En definitiva, según los autores citados hay muchas oportunidades para aumentar y desarrollar mejoras en relación a las tesis electrónicas. Pueden estar involucrados diferentes aspectos, habiendo un margen de mejora considerable que puede llevar a conseguir una mejor calidad de los contenidos y los de servicios. También se pueden aprovechar estos aspectos en el trabajo cooperativo relacionado con movimientos como el del acceso abierto, los repositorios y las bibliotecas digitales. Se podrían aplicar además algunas contribuciones técnicas que ayuden a la búsqueda, la navegación, la vinculación y otros servicios.

Capítulo I.6. Estudios de impacto de la investigación en acceso abierto. Visibilidad, uso y citación de las tesis doctorales

En este capítulo se realiza un breve estado de la cuestión de los beneficios del acceso abierto relativos a la visibilidad y el impacto a través de la literatura existente sobre los estudios de impacto de la investigación puesta en acceso abierto a través de repositorios, es decir la vía verde del acceso abierto. Se hará hincapié sobre todo en la visibilidad, uso y citación que alcanzan las tesis doctorales depositadas en repositorios.

Prácticamente todas las bibliotecas académicas del mundo apoyan el acceso abierto verde (*Green OA*) mediante la creación y mantenimiento de repositorios institucionales que difunden la producción académica y de investigación de la universidad, bajo la forma de tesis, actas de congresos, publicaciones universitarias, materiales históricos y patrimoniales, datos de investigación y artículos de profesores e investigadores. El profesorado que pone en acceso abierto su investigación se beneficia por una parte de un aumento exponencial de lectores, es decir, de un aumento de visibilidad y de uso, y por otro lado del impacto de dicha investigación, esto es, de una ventaja de citación que ha sido demostrada en numerosos estudios.

Por desgracia, estos beneficios por sí solos no son el resultado de la vía verde del acceso abierto, es decir, del autoarchivo en repositorios realizado por los autores sobre sus artículos. En 2005 *The National Institutes of Health* (NIH) estableció su primera política de acceso público para poner a disposición del público gratuitamente los manuscritos finales, revisados por pares y resultantes de la investigación financiada por el NIH, a través de PubMed Central y dentro de los 12 meses después de la publicación en una revista científica. Como esta medida política solo animaba al depósito de artículos, se mantuvo en gran medida inadvertida hasta que en 2008 se promulgó una política más fuerte que requería el depósito. La tasa de depósito se situaba en el

7% para el período de 2005 a 2007 aumentando al 75% en marzo de 2012 (Executive Office of the President National Science and Technology Council, 2012).

Las universidades han estado dispuestas a rendir cuentas con el depósito de artículos en los repositorios institucionales y se han esforzado por aumentar las tasas de depósito de los artículos entre sus profesores. A pesar de que existen políticas de acceso abierto aprobadas en universidades de todo el mundo y a menudo se requiere el depósito de artículos en los respectivos repositorios institucionales, estas políticas generalmente no se hacen cumplir, variando así ampliamente las tasas de depósito en las diferentes instituciones.

Con todo esto, se quiere expresar que lamentablemente aunque los beneficios del acceso abierto verde se demuestran continuamente, cada vez que se realiza un estudio sobre el aumento de visibilidad y de uso de la de la investigación abierta, así como la ventaja de citación sobre la misma, algunos autores son todavía reacios a autoarchivar voluntariamente su investigación y solo lo hacen cuando existe una sólida y fuerte política de mandato de depósito en su institución.

I.6.1. Estudios de impacto de la investigación en acceso abierto

SPARC Europe (SPARC Europe, 2015) mantiene en su web una lista de estudios y resultados sobre la ventaja de citación del acceso abierto. Para realizar el estado de la cuestión sobre este tema, se toma como referencia la lista SPARC, eliminando algunos estudios y añadiendo otros que se consideran también interesantes. Toda la literatura referenciada sobre el tema se obtuvo a través de la revisión sistemática de la literatura llevada a cabo para la elaboración de la bibliografía de esta tesis.

En la Tabla 31 se indica en la primera columna las referencias de los estudios, en formato de cita APA e importadas de la bibliografía seleccionada por la doctoranda en *Mendeley*, y en la

segunda columna se indican el análisis y resultados, expresando si los autores consideran que existe ventaja de citación.

Tabla 31

Estudios sobre la ventaja de citación del acceso abierto

#	ESTUDIO	¿Existe ventaja de citación?
1	Atchison, A., & Bull, J. (2015). Will Open Access Get Me Cited? An Analysis of the Efficacy of Open Access Publishing in Political Science. <i>PS: Political Science & Politics</i> . Retrieved from http://scholar.valpo.edu/ccls_fac_pub/21	<ul style="list-style-type: none"> - Se compara las tasas de citación medias entre artículos OA y artículos de suscripción - Se concluye que existe ventaja de citación: "la publicación OA resulta en clara ventaja de citación"
2	Donovan, J. M., Watson, C. A., & Osborne, C. (2014). The Open Access Advantage for American Law Reviews. <i>SSRN Electronic Journal</i> . doi:10.2139/ssrn.2506913	<ul style="list-style-type: none"> - Se comparan las citas de artículos OA por año con los artículos publicados en la mismas revistas pero que no están en OA - Hay ventaja. Los artículos OA acumulan 49% más citas por año que los no OA. Para los artículos publicados en revistas de nivel superior la ventaja de cita fue menor
3	Frisch, N. K., Nathan, R., Ahmed, Y. K., Shidham, V. B., & Ahmed, Y. K. (2014). Authors attain comparable or slightly higher rates of citation publishing in an open access journal (CytoJournal) compared to traditional cytopathology journals - A five year (2007-2011) experience. <i>CytoJournal</i> , 11(1), 10. doi:10.4103/1742-6413.131739	<ul style="list-style-type: none"> - Se comparan citas con artículos de los mismos autores publicados en revistas OA y no OA. El tiempo transcurrido para que se acumulen las citas: 5 años - Las publicaciones en revistas de acceso abierto recibieron significativamente más citas en promedio que los de las revistas no OA ($p < 0,05$). Sin embargo, con los resúmenes de congresos excluidos del data set, la diferencia no fue significativa ($p = 0,57$)
4	Koler-Povh, T., Južnič, P., & Turk, G. (2014). Impact of open access on citation of scholarly publications in the field of civil engineering. <i>Scientometrics</i> , 98(2), 1033–1045. doi:10.1007/s11192-013-1101-x	<ul style="list-style-type: none"> - Se compara el promedio de citas por artículo entre artículos OA y no OA publicados en las mismas revistas. El tiempo transcurrido para que las citas se acumulen: 5 años - Existe ventaja de citación. En general, el promedio de citas por artículo fue significativamente mayor para los artículos OA que no para los artículos no OA. Los artículos OA forman el 22% de la muestra, pero recibieron el 29% del total de citas de la muestra. Este patrón se producía en todas las revistas de mayor impacto estudiadas y para algunas revistas de impacto medio, pero no para las revistas de menor impacto
5	Kullman L (2014) The Effect of Open Access on Citation Rates of Self-archived Articles at Chalmers. In IATUL 2014–35th Annual Conference–Aalto University, Espoo, Finland, 2-5 June 2014.	<ul style="list-style-type: none"> - Se compara las puntuaciones medias de las referencias normalizadas (MNCS número de citas normalizadas por campo y año de publicación) con los artículos listados con texto completo con datos bibliográficos solo en el repositorio universitario - Hay ventaja de citación. Los artículos a texto completo tenían una tasa de citas del 22% más alta que los que solo tenían datos bibliográficos, y esta diferencia fue significativa

#	ESTUDIO	¿Existe ventaja de citación?
6	McCabe MJ and Snyder CM (2014) Identifying the effect of Open Access on citations using a panel of science journals. <i>Economic Inquiry</i> , 52(4): 1284–1300.	<ul style="list-style-type: none"> - Se compara el número de citas de los volúmenes completos de revistas OA con los volúmenes de revistas que son parcialmente OA y no OA. El tiempo transcurrido para que las citas se acumulen: máximo 10 años - En general hay ventaja. Se observa un aumento del 8% citación para todo el conjunto de datos. Sin embargo, depende del impacto de la revista. Las citas de revistas de alto nivel se incrementaron significativamente con el OA, pero las citas de revistas de bajo nivel disminuyeron significativamente con el OA
7	Mueller-Langer F & Watt R (2014) The Hybrid Open Access Citation Advantage: How Many More Cites is a \$3,000 Fee Buying You? Max Planck Institute for Innovation & Competition Research Paper No. 14-02.	<ul style="list-style-type: none"> -Se compara el número de citas entre artículos publicados OA y no OA en las mismas revistas. El tiempo transcurrido para que las citas se acumulen: 1-12 años - No hay ventaja; aunque hubo significativamente más citas para los artículos OA que para los no OA, la diferencia se eliminó mediante el control de calidad de la institución y por la existencia de un preprint OA
8	Laakso, M., & Bjork, B.C. (2013). Delayed Open Access: An Overlooked High-Impact Category of Openly Available Scientific Literature. <i>Journal of the American Society for Information Science and Technology</i> , 64(7), 1323–1329.	<ul style="list-style-type: none"> - Se comparan los factores de impacto 2010 entre los factores de impacto de revistas OA atrasadas (las que permiten OA verde), revistas de acceso abierto dorado y revistas de suscripción. El tiempo transcurrido para que las citas se acumulen: los factores de impacto están basados en los 5 años de citas - Hay ventaja de citación. Como promedio las revistas OA verdes tienen tasas de citación el doble de altas que las revistas que son solo de suscripción, y tres veces más altas que las revistas de OA inmediato
9	Piwowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. <i>PeerJ</i> , 1, e175. doi:10.7717/peerj.175	<ul style="list-style-type: none"> - Se comparan el número de citaciones para los paper que han puesto disponibles en acceso abierto sus <i>datasets</i> a través de sus repositorios con aquellos que no han publicado sus datos. El estado OA de los paper originales fue controlado como una covariante.El tiempo transcurrido para que las citas se acumulen fue de 2 a 10 años. - Hay ventaja de citación. Los papers con conjuntos de datos de libre acceso, recibieron una media del 9% (intervalo de confianza del 95%: 5% a 13%) más citas que los datos no publicados
10	Riera, M., & Aibar, E. (2013). ¿Favorece la publicación en abierto el impacto de los artículos científicos? Un estudio empírico en el ámbito de la medicina intensiva. <i>Medicina Intensiva</i> , 37(4), 232–240. doi:10.1016/j.medin.2012.04.002	<ul style="list-style-type: none"> - Se compara en bruto el número de citas entre artículos OA y artículos no OA artículos. El tiempo transcurrido para que las citas se acumulen: 2-3 años. - Hay ventaja de citación; los artículos OA reciben mensualmente más citaciones sobre el promedio de los artículos que no son de acceso abierto. Aunque no hubo una diferencia significativa en el número de citas en bruto entre todos los artículos OA y no OA el acceso abierto

#	ESTUDIO	¿Existe ventaja de citación?
		y artículos de acceso no abiertas, había una diferencia significativa en el número de citas entre el acceso abierto más citado y artículos de acceso no abiertos
11	Solomon, D. J., Laakso, M., & Björk, B.-C. (2013). A longitudinal comparison of citation rates and growth among open access journals. <i>Journal of Informetrics</i> , 7(3), 642–650. doi:10.1016/j.joi.2013.03.008	<ul style="list-style-type: none"> - Se compara SNIP2 (Source-Normalized Impact per Paper –a field-adjusted citation impact measure) para las revistas OA financiadas por APCs, revistas OA financiadas por otros medios, y revistas por suscripción - No hay ventaja de citación, los promedios SNIP2 para las revistas OA financiadas por APC fueron más o menos igual o inferiores a las revistas por suscripción, mientras que los de otras publicaciones OA eran aún más bajos
12	Vanclay, J. K. (2013). Factors affecting citation rates in environmental science. <i>Journal of Informetrics</i> , 7(2), 265–271. doi:10.1016/j.joi.2012.11.009	<ul style="list-style-type: none"> - Se codificó el estado de las publicaciones como una variable ternaria: 0= revista por suscripción, 1= OA verde, 2= OA dorado. Se utilizó la prueba de correlación de Pearson entre el estado OA y las citas por año. El tiempo transcurrido para que se acumulen las citas: 5-6 años - No hay ventaja de citación; no hubo una correlación significativa entre el estado OA y las tasas de citación
13	Björk, B.-C., & Solomon, D. (2012). Open access versus subscription journals: a comparison of scientific impact. <i>BMC Medicine</i> , 10(1), 73. Retrieved from http://www.biomedcentral.com/1741-7015/10/73	<ul style="list-style-type: none"> - Se comparan 2 años de los factores de impacto de revistas OA y de las revistas de suscripción, se controlaron la edad revista, el país de la editorial, la disciplina y el modelo de negocio del OA - No hay ventaja de citación; el promedio de las tasas de citación fueron alrededor del 30% más altas para las revistas por suscripción que para las revistas OA. Después de controlar la disciplina, la edad de la revista, y la ubicación de la editorial, la diferencia desapareció para la mayoría de las subcategorías
14	Mahesh G. (2012) Open access and impact factors. <i>Current Science</i> , 103(6), 610.	<ul style="list-style-type: none"> - Se analiza el cambio en factores de impacto de las revistas NISCAIR en los años 2008 y 2009 cuando se convirtieron en OA - Hay ventaja de citación. La casi totalidad de las 17 revistas aumentaron sus índices de impacto en los años inmediatamente posteriores al cambio a OA
15	Xia, J., & Nakanishi, K. (2012). Self-selection and the citation advantage of open access articles. <i>Online Information Review</i> , 36(1), 40–51	<ul style="list-style-type: none"> - Se comparan citas entre artículos OA y no OA por separado de revistas situadas arriba y abajo en el <i>ranking</i> - Hay ventaja de citación en todas ellas. Los artículos de las revistas situadas más abajo en el <i>ranking</i> también han tenido una tasa de citación más alta. Sin embargo, los artículos situados en lo alto del <i>ranking</i> no han tenido una tasa de citación si se hacen OA

#	ESTUDIO	¿Existe ventaja de citación?
16	Davis, P. M. (2011). Open access, readership, citations: a randomized controlled trial of scientific journal publishing. <i>The FASEB Journal</i> , 25(7), 2129–2134. doi:10.1096/fj.11-183988	<ul style="list-style-type: none"> - Se comparan artículos descargados 12 meses después de la publicación y citados después de 3 años entre artículos OA dorado y artículos no OA/OA verde - No hay ventaja de citación. Artículos OA fueron citados más frecuentemente que los artículos de acceso por suscripción. Sin embargo, las versiones HTML a texto completo y las versiones PDF de artículos OA fueron descargadas significativamente más frecuentemente durante los primeros 12 meses
17	Donovan, J. M., & Watson, C. A. (2011). Citation advantage of open access legal scholarship. <i>Law Library Journal</i> , 103(4), 553–573	<ul style="list-style-type: none"> - Se comparan la citación anual de artículos OA y de artículos no OA dentro de las mismas revistas. El tiempo transcurrido para que las citas se acumulen es de hasta 18 años - Hay ventaja de citación. Los artículos OA recibieron, como promedio, un 58% más citas que los artículos no-OA de edad y calidad similar en la misma revista
18	Davis, P. M. (2010). Does open access lead to increased readership and citations? A randomized controlled trial of articles published in aps journals. <i>Physiologist</i> , 53(6), 197–201	<ul style="list-style-type: none"> - Se comparan descargas y citas entre artículos hechos OA inmediatamente de su publicación y aquellos artículos hechos OA 12 meses después de su publicación en revistas solo de suscripción. El tiempo transcurrido para que las citas se acumulen es de 3 años - No hay ventaja de citación. Los artículos OA dorado no fueron más propensos a ser citados durante el primer año después de su publicación y no fueron más frecuentemente citados durante los primeros tres años que los artículos OA verde. Sin embargo, los artículos OA dorado recibieron significativamente más descargas del texto completo y PDF que los artículos OA verde durante el primer año, que eran solo suscripción durante este periodo
19	Gargouri, Y., Hajjem, C., Larivière, V., Gingras, Y., Carr, L., Brody, T., & Harnad, S. (2010). Self-selected or mandated, open access increases citation impact for higher quality research. <i>PLoS ONE</i> , 5. doi:10.1371/journal.pone.0013636	<ul style="list-style-type: none"> - Se comparan artículos OA con artículos que no se hicieron OA publicados en el mismo número de de las mismas revistas. Se hace un análisis por regresión logística. El tiempo transcurrido para que se acumulen las citas fue de 2 a 6 años - Hay ventaja de citación y el tamaño de la ventaja de citación varía según la disciplina. No fue significativa la reducción de la ventaja OA encontrada con o sin mandato. La ventaja encontrada fue independiente del Journal Impact Factor, el tiempo transcurrido desde la publicación, el número de referencias en el artículo y el número de coautores. Se vió que el aumento de citaciones para los artículos OA fue más fuerte los artículos más citados. Los artículos procedentes del del mandato institucional incrementaron las citas en un rango medio a alto. Los artículos revisados incrementaron las citas y el efecto es más notorio en los artículos con un grado de citación alto

#	ESTUDIO	¿Existe ventaja de citación?
20	Snijder, R. (2010). The profits of free books: an experiment to measure the impact of open access publishing. <i>Learned Publishing</i> , 23(4), 293–301. doi:10.1087/20100403	<ul style="list-style-type: none"> - Se analizó el impacto de libros. Se crearon 4 conjuntos con 100 libros cada uno: libros no OA, libros disponibles libremente a través de Google Book Search, libros disponibles libremente a través del repositorio AUP (Amsterdam University Press), y libros disponibles libremente - No hubo ventaja de citación. No se encontró una relación directa con el estatus OA y las tasas de citación. Sin embargo contrariamente a lo esperado, no hubo disminución en las ventas de los libros que se encontraban en acceso abierto
21	Xia, J., Myers, R. L., & Wilhoite, S. K. (2011). Multiple open access availability and citation impact. <i>Journal of Information Science</i> , 37(1), 19–28. doi:10.1177/0165551510389358	<ul style="list-style-type: none"> - Se compara el número de citaciones entre artículos no OA, de los que tenían una sola copia OA, y de los que tenían disponibles en línea varias copias. El tiempo transcurrido para que se acumulen las citas es de 4 años - Hay ventaja de citación. Los artículos OA recibieron un promedio de 2,13 citas más que los artículos no OA. Además los artículos OA recibieron una media de 2,348 citas más por copia disponible en línea
22	Bernius, S., & Hanauske, M. (2009). Open Access to Scientific Literature - Increasing Citations as an Incentive for Authors to Make Their Publications Freely Accessible. In 42nd Hawaii International Conference on System Sciences (HICSS '09) (pp. 1–9). IEEE. doi:10.1109/HICSS.2009.335	<ul style="list-style-type: none"> - Los autores desarrollaron una simulación por ordenador de redes de citación, utilizando una serie de supuestos para los parámetros, con el fin de modelar el comportamiento de citación y los resultados. La metodología intentaba incluir tanto la casualidad como el efecto del modelo. - Hay ventaja de citación. Los resultados de la simulación apoyan los datos empíricos sobre el aumento de las citas de los artículos que se publican bajo el paradigma de acceso abierto
23	Evans, J. A., & Reimer, J. (2009). Open access and global participation in science. <i>Science (New York, N.Y.)</i> , 323(5917), 1025. Retrieved from http://www.sciencemag.org/cgi/content/abstract/323/5917/1025	<ul style="list-style-type: none"> - Se compararon las citas de revistas OA y revistas no OA. - Hay ventaja de citación. El incremento en citación con OA es de un 8% para los artículos publicados recientemente, y dos veces más para las citas que provenían de países en vía de desarrollo. El estudio también encontró que hay un incremento en las citaciones cuando los artículos se ponen a disposición libremente por primera vez al finalizar un período de embargo
24	Frandsen, T. F. (2009). The effects of open access on unpublished documents: A case study of economics working papers. <i>Journal of Informetrics</i> , 3(2), 124–133. doi:10.1016/j.joi.2008.12.002	<ul style="list-style-type: none"> - Se analizaron los efectos del acceso abierto sobre trabajos no publicados como los documentos de trabajo. El tiempo transcurrido para que citas se acumulen: las citas de los documentos de trabajo se midieron durante un período de diez años - No se encontró ninguna tendencia clara de un aumento en el impacto durante los 10 años de disponibilidad en acceso abierto de los documentos de trabajo. A la inversa, artículos de economía, en revistas de alto impacto, no muestran una tendencia clara de impacto de citación a aumentar en un período de diez años.

#	ESTUDIO	¿Existe ventaja de citación?
		Por consiguiente, el autor deduce que no hay ninguna ventaja de citación OA
25	Norris, M., Oppenheim, C., & Rowland, F. (2008). The citation advantage of open-access articles. <i>Journal of the American Society for Information Science & Technology</i> , 59(12)	<ul style="list-style-type: none"> - Se comparan las citas recibidas por los artículos de revistas con aquellos que también se pusieron en OA. El tiempo transcurrido para que se acumulen las citas: 2 a 4 años - Hay ventaja de citación. La media de citas para los artículos OA fue de 9,04 y para los artículos no OA de 5,76. El tamaño de la ventaja varió con la disciplina: la sociología mostró la ventaja más alta y la ecología la más baja. La correlación con el número de autor y el factor de impacto de la revista es débil o inexistente
26	Cheng, W., & Ren, S. (2008). Evolution of open access publishing in Chinese scientific journals. <i>Learned Publishing</i> , 21(2), 140+. doi:10.1087/095315108X288884	<ul style="list-style-type: none"> - Se examinaron citas, índice de inmediatez y factor de impacto de revistas de acceso abierto y revistas no OA en las 4 áreas en el conjunto de revistas universitarias producidas. Tiempo transcurrido para que las citas se acumulen: 2 años (enfoque basado en la metodología del Journal Citation Index para calcular el JIF) - Hay ventaja de citación: hubo un aumento de aproximadamente el doble en las citas de revistas de acceso abierto
27	Davis, P. M., Lewenstein, B. V, Simon, D. H., Booth, J. G., & Connolly, M. J. L. (2008). Open access publishing, article downloads, and citations: randomised controlled trial. <i>BMJ</i> , 337(jul31 1), a568–a568. doi:10.1136/bmj.a568	<ul style="list-style-type: none"> - Se tomó un conjunto de artículos de los 4 números de 11 revistas OA de American Physiological Society (elegidas al azar) y se dejó el resto no OA. Todas estas revistas hacen OA su contenido al cabo de 12 meses después de publicarse. El tiempo transcurrido para que las citas se acumulen: 12 meses - No se encontró ventaja de citación. Se vió que las descargas del texto completo aumentaron un 89% con el acceso abierto y que las descargas de PDFs aumentaron un 42%, sin embargo el acceso abierto redujo las citas en un 5%
28	Frandsen, T. F. (2009). The integration of open access journals in the scholarly communication system: Three science fields. <i>INFORMATION PROCESSING & MANAGEMENT</i> , 45(1), 131–141. doi:10.1016/j.ipm.2008.06.001	<ul style="list-style-type: none"> - Se compararon las citas de revistas de suscripción con revistas que son OA y con aquellas que no son OA. Se controló la autocita, el JIF y la dependencia de muestreo. El tiempo transcurrido para que las citas se acumulen: 2 años - No hay ventaja de citación. El efecto del OA sobre las citas fue neutral en el área de la biología y en el área de las matemáticas y negativo en farmacia y farmacología
29	Gaulé, P., & Maystre, N. (2011). Getting cited: Does open access help? <i>Research Policy</i> , 40(10), 1332–1338. doi:10.1016/j.respol.2011.05.025	<ul style="list-style-type: none"> - Se compararon las citas acumuladas a lo largo del tiempo de los artículos de una revista que se hicieron OA como resultado del pago de los autores (APC) con los artículos que no eran OA. El tiempo transcurrido para que las citas se acumulen: 2 años.- Hay ventaja de citación pero la diferencia no fue estadísticamente significativa

#	ESTUDIO	¿Existe ventaja de citación?
30	Norris, M., Oppenheim, C., & Rowland, F. (2008). The citation advantage of open-access articles. <i>Journal of the American Society for Information Science & Technology</i> , 59(12)	<ul style="list-style-type: none"> - Se compararon las citas de artículos OA y no OA en las disciplinas de ciencias de la información y tecnología, usando artículos publicados en la misma revista. El tiempo transcurrido para que las citas se acumulen: de 2 a 4 años - Hay ventaja de citación. La media para los artículos OA fue de 9,04 y para los artículos no OA 5,76
31	Piwowar HA, Day RS & Fridsma DB (2007) Sharing Detailed Research Data Is Associated with Increased Citation Rate. <i>PLoS ONE</i> , March 21, 2007	<ul style="list-style-type: none"> - Se buscó la disponibilidad de datos abiertos soportando los artículos para ver si estaba relacionada con el impacto de citación. El tiempo transcurrido para que las citas se acumulen: 4 a 7 años - Los conjuntos de datos disponibles para el público (datos abiertos) se asociaron significativamente con un aumento del 69% en las citas a los artículos que acompañan a los datos. Esta correlación es independiente del factor de impacto, el país de los autores y el tiempo transcurrido desde la publicación
32	Davis, P. M., & Fromerth, M. J. (2007). Does the arXiv lead to higher citations and reduced publisher downloads for mathematics articles? <i>Scientometrics</i> , 71(2), 203–215. <i>Digital Libraries; Information Retrieval; History and Overview</i> . doi:10.1007/s11192-007-1661-8	<ul style="list-style-type: none"> - Se comparan las citas de artículos en revistas con citas de artículos publicados en esas mismas revistas pero depositados en arXiv. El tiempo transcurrido para que las citas se acumulen: de 6 meses a 8 años - Hay ventaja de citación. El aumento medio es del 35% (el número de citas a cada artículo aumento entre un 0,8 a un 2,1; produciéndose un aumento medio de 1,1 que corresponde a un aumento de 35% del total)
33	Eysenbach G (2006) Citation Advantage of Open Access Articles. <i>PLoS Biology</i> , 4(5), May 2006	<ul style="list-style-type: none"> - Se comparan las citas de artículos OA y no OA publicados por PNAS durante un período de 6 meses. El tiempo transcurrido para que las citas se acumulen: 18 meses - Hay ventaja de citación. Los artículos OA son 2,1 veces más propensos a ser citados en los primeros 4-10 meses después de su publicación
34	Moed HF (2006) The effect of 'Open Access' upon citation impact: An analysis of ArXiv's Condensed Matter Section. <i>Journal of the American Society for Information Science and Technology</i> , 58(13): 2145-2156	<ul style="list-style-type: none"> - Se comparan las citas a artículos depositados en arXiv con aquellos artículos en las mismas revistas que no se pusieron disponibles a través de arXiv. El tiempo transcurrido para que las citas se acumulen: 12 meses a 14 años - El estudio no fue diseñado específicamente para explorar si el OA conlleva impacto de citación extra sino para probar los efectos en general del OA frente al <i>Early Access</i> y frente al sesgo de selección
35	Zhang Y (2006) The Effect of Open Access on Citation Impact: A Comparison Study Based on Web Citation Analysis. <i>Libri</i> , 56(3): 133-199	<ul style="list-style-type: none"> - Se comparan las citas de artículos en dos revistas. Se recuperaron citas en la web usando Google and Yahoo. El tiempo transcurrido para que las citas se acumulen: 4-5 años - Hay ventaja de citación. Como promedio los artículos OA recibieron dos veces más citas que

#	ESTUDIO	¿Existe ventaja de citación?
36	Hajjem C, Harnad S & Gingras Y (2005) Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact. IEEE Data Engineering Bulletin, 28(4), December 2005	<p>aquellos que no son OA. El estudio también identificó que el mayor aumento de citas procedía de documentos no autorizados (las otras dos categorías fueron los documentos académicos y documentos docentes). El estudio también encontró un aumento de citación de los autores en los países en desarrollo</p> <ul style="list-style-type: none"> - Se rastreó la web mediante un robot para buscar artículos en acceso libre y se emparejaban con artículos del mismo número de la misma revista (utilizando WOS) que solo estaban disponibles en la revista por suscripción. El tiempo transcurrido para que las citas se acumulen: 6 a 12 meses - Hay ventaja de citación. OA produce un aumento de citación entre un 36% y un 172%
37	Antelman K (2004) Do Open-Access Articles Have a Greater Research Impact? College and Research Libraries, 65(5): 372-382, September 2004	<ul style="list-style-type: none"> - Se compararon citas a partir de 2-3 años de artículos de diez revistas líderes en cada disciplina con citas de artículos de esas mismas revistas que estaban disponibles libremente en la Web. El tiempo transcurrido para que las citas se acumulen: a partir de 2 años - Hay ventaja de citación. Aumentan las citas: 91% en matemáticas, ingeniería eléctrica 51%, ciencia política 86%, filosofía 45%
38	Harnad S & Brody T (2004) Comparing the impact of open access (OA) vs non-OA articles in the same journals. D-Lib Magazine, 10(6), June. (reporting results from this study: http://opcit.eprints.org/feb19oa/brody-impact.pdf)	<ul style="list-style-type: none"> - Se comparan los artículos en los campos de la física que están disponibles abiertamente en arXiv con aquellos en los mismos números de las mismas revistas que no son OA. El tiempo transcurrido para que las citas se acumulen: de 2 a 12 años - Hay ventaja de citación. El incremento es del 250% al 500%
39	Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., Henneken, E., & Murray, S. S. (2005). The effect of use and access on citations. Information Processing & Management, 41(6), 1395–1402. doi:10.1016/j.ipm.2005.03.010	<ul style="list-style-type: none"> - Se utilizaron citas de artículos de siete revistas de astrofísica a otros artículos en esas revistas. Se comparon las citas a artículos publicados solamente en las revistas con citas a los artículos publicados en esas revistas pero depositados en arXiv. El tiempo transcurrido para que se acumulen las citas: seguimiento de las citas desde la publicación hasta 20 años - Hay ventaja de citación. Se confirmo el sesgo de selección y la ventaja de lo más reciente. No se encontró una ventaja al OA general pero esto se explica porque todos los investigadores de la astronomía tienen acceso de cualquier modo a toda la literatura de astronomía

#	ESTUDIO	¿Existe ventaja de citación?
40	Lawrence S (2001) Free <i>online</i> availability substantially increases a paper's impact. Nature, 31 May 2001.	- Se tuvo en cuenta la citación marcada contra la disponibilidad libre en línea de los artículos y se demostró que los artículos disponibles libremente suelen tener más altas tasas de citación. El tiempo transcurrido para que se acumulen las citas: de 1 a 11 años - Hay ventaja de citación. Se encontró que el incremento medio de citas de artículos OA era de 157%. Para las publicaciones del tipo de las conferencias la media de incremento fue de 284%

Nota. Fuente: Elaboración propia

I.6.2. Visibilidad, uso y citación de las tesis doctorales

Como se ha podido observar a través de la revisión de la literatura realizada en este capítulo los estudios realizados sobre el impacto del acceso abierto determinan en su mayoría que existe ventaja de citación sobre los artículos puestos en acceso abierto, es decir, el número de citas recibidas por los artículos en acceso abierto es mayor que el número de citas recibidas por los artículos que no se encuentran en acceso abierto.

Hajjem, Harnad y Gingras (2005) realizaron un estudio basado en un gran conjunto de datos de más de un millón de artículos publicados en el periodo de 1992 a 2003 en diez áreas científicas, y en ese estudio llegaron a la conclusión de que los documentos en acceso abierto han recibido sistemáticamente más citas que los que no estaban en acceso abierto.

Harnad y Brody (2004) encontraron que la proporción anual de documentos de acceso abierto ha ido creciendo significativamente más rápidamente que la de los artículos no-OA, mediante la comparación de los dos tipos de artículos en la misma revista y el mismo año.

Es evidente que el acceso abierto proporciona visibilidad a las tesis depositadas en repositorios. Pero ¿presentan las tesis alojadas en los repositorios de acceso abierto una ventaja de citación como los artículos científicos?

El objetivo de esta tesis doctoral es demostrar que los beneficios del acceso abierto recaen también sobre la literatura gris científica en acceso abierto, de la que son un claro exponente las e-tesis. Aunque no son tan numerosos los estudios sobre visibilidad e impacto de las e-tesis, los que se han realizado hasta el momento demuestran que el acceso abierto es la mejor forma de comunicar las e-tesis, no solo por la contribución que se hace a la sociedad en general al publicar libremente los resultados de la investigación científica, sino también porque tanto autores como instituciones se benefician de las bondades del acceso abierto: accesibilidad, difusión, visibilidad e impacto.

Junto con otros documentos y artículos, la literatura gris contribuye al éxito de los repositorios institucionales. Su naturaleza no comercial y alternativa sitúa a la literatura gris muy próxima a la cultura impulsada por la comunidad de los archivos abiertos. Los repositorios de acceso abierto han formado una parte importante de la implementación del mismo desde el comienzo del movimiento de acceso abierto (Björk et al., 2014).

Actualmente, las tesis doctorales se encuentran en su mayoría depositadas en acceso abierto a través de los repositorios institucionales. Uno de los beneficios del depósito en los RIs es la disponibilidad para todo el mundo. Potencialmente, esto beneficia tanto a la comunidad académica al hacer accesible la investigación actual, y a los estudiantes al proporcionarles un espacio desde donde hacer accesibles públicamente sus trabajos para que puedan ser recuperados y utilizados.

La inclusión de las tesis en el repositorio incrementa las estadísticas de uso del mismo y permite que la investigación original realizada por los investigadores principiantes sea visible, accesible y capaz de ser utilizada por la comunidad científica en general. Las estadísticas indican que las tesis doctorales que están disponibles de este modo son las más utilizadas y citadas en la investigación en curso (Stanton & Liew, 2011; Troman et al., 2007).

Larivière, Zuccala y Archambault (2008) estudiaron el impacto de las tesis doctorales y los resultados de su estudio fueron paradójicos, ya que mientras que las tesis eran cada vez más accesibles en formato electrónico para los académicos, su impacto científico en general parecía disminuir. Esto no significa necesariamente que una tesis doctoral sea una fuente pobre de información académica. La razón puede ser que los investigadores prefieren citar los artículos y libros publicados derivados de la investigación, en lugar de las tesis originales. En algunos casos, puede ser que la disponibilidad de las tesis como fuentes de información se haya pasado por alto, dada la creciente cantidad de otros tipos de publicaciones (artículos de revistas, informes de investigación, etc.) que están disponibles en la Web (Vincent Larivière et al., 2008).

Por otra parte, algunos estudios dirigen su centro de atención hacia los elementos de interoperabilidad como una de las causas que incrementan la diseminación y visibilidad de las tesis doctorales en los repositorios. Para los investigadores y las instituciones es esencial conseguir visibilidad, citación e impacto de su trabajo. Por ello es muy importante que los repositorios aseguren la interoperabilidad y la preservación digital con el fin de conseguir difusión, visibilidad, accesibilidad y uso a largo plazo de su contenido digital (Friend, 2011; Hakimjavadi & Masrek, 2013; Schöpfel, 2013).

Asimismo, las estadísticas de descarga de las tesis doctorales desde el propio repositorio sugieren un uso extenso. Un estudio de Stone y Lowe (2014) explora qué significan esas estadísticas de descarga, y cómo las tesis están siendo utilizadas e integradas más ampliamente en el conjunto de la investigación científica. Los autores se preguntan por un lado, si las tesis se utilizan y se citan en fuentes reconocidas y en revistas científicas, y por otro lado dónde acaba la investigación de los doctorandos. Esta investigación muestra como ejemplo un caso de una popular tesis depositada en el RI en enero de 2013 que fue descargada 3.195 veces y citada tres veces en diciembre de 2013. Los autores se plantean si están relacionadas las altas tasas de descargas con las citas de las tesis, qué significa este acceso sobre el valor de las tesis. A través

de este estudio realizado con *Google Scholar* los autores demuestran que las tesis son citadas en revistas reconocidas, revistas revisadas por pares, aunque no al ritmo que se cita el material científico. Stone y Lowe (2014) afirman que el depósito de las tesis en el repositorio es beneficioso para los autores de las tesis ya que estas se ponen a disposición de toda la comunidad científica brindando a los estudiantes la oportunidad de difundir su trabajo, aumentando su visibilidad, el uso de su trabajo y la citación del mismo.

Otro aspecto tratado actualmente muy a menudo por los autores es el de la citación en las tesis doctorales, aspecto que indica el comportamiento de citación de los nuevos investigadores. Para Varshney (2012) los avances tecnológicos a lo largo del tiempo han hecho posible el aumento de productividad de los investigadores. El autor presenta un estudio longitudinal del comportamiento de la citación en las tesis doctorales del Departamento de Ingeniería Eléctrica e Informática del *Massachusetts Institute of Technology* (MIT), durante el período 2004-2011, hallando que el número de referencias citadas ha aumentado a lo largo del tiempo; el autor indica que al mismo tiempo, ha habido una disminución en la duración del programa de doctorado y una constante relativa de la cultura del departamento; esto sugiere que los estudiantes son más productivos a la hora de hacer frente a una mayor carga de conocimiento y parecen codificar la literatura anterior como memoria transactiva, como lo demuestra el mayor uso de la literatura más antigua. Varshney (2012) cita alguna de las teorías que argumentan que el número de referencias citadas en las tesis doctorales ha aumentado por la introducción de las nuevas tecnologías como *Google Scholar*, ya que se facilita un mayor uso de la literatura como memoria transactiva. Los hallazgos cuantitativos de Varshney (2012) también sugieren también un “efecto *Google*” que indica un cambio cognitivo en la investigación de los estudiantes. Cuando los investigadores recuerdan dónde está almacenada la información en lugar de la información en sí misma, la naturaleza de su investigación es posible que cambie. En último término, este artículo viene a demostrar que la información difundida a través de Internet y recuperable a través de *Google*, por ejemplo,

aumenta las posibilidades de ser citada, lo cual es una oportunidad para las tesis doctorales depositadas en repositorios de acceso abierto.

Otra forma de explorar la visibilidad y el descubrimiento de las tesis doctorales electrónicas es a través de los datos extraídos y analizados de *Google Analytics* (Cheng, Cheung, & Kot, 2014). Según el estudio de estos autores los datos generados desde el repositorio digital de tesis doctorales de la Universidad de Lingnan y de *Google Analytics*, entre abril de 2012 y diciembre de 2013, pueden utilizarse para entender mejor cómo Internet permite el descubrimiento de las tesis doctorales electrónicas aumentando el uso de las mismas.

Por otra parte, existen algunos análisis bibliométricos, algunos ya referenciados anteriormente en el capítulo I.4., sobre el uso e impacto de la literatura gris que, por extensión, podrían aplicarse al análisis de las tesis doctorales (Cesare et al., 2008; Cesare & Ruggieri, 2006).

Cesare y Ruggieri (2006) aplicaron una metodología de análisis de citación de LG para descubrir su impacto en diferentes disciplinas a través del uso de indicadores bibliométricos. Las autoras proponen una metodología para analizar el uso de la literatura gris en las referencias bibliográficas de los artículos de revistas. Con el fin de verificar en qué medida la literatura gris contribuye al trabajo científico, se utilizaron indicadores bibliométricos similares a los utilizados para la literatura convencional, tales como la frecuencia de uso y la frecuencia de cita de la LG. Las autoras describen un indicador bibliométrico específico llamado “intensidad de citación de LG”, que resulta útil para medir el impacto de la literatura gris, es decir, la frecuencia media de referencias de LG por artículo que contenga LG. Se tiene en cuenta el uso potencial del factor de impacto como un indicador de la variabilidad en el uso de LG. Los resultados del análisis confirman que la literatura gris es una importante fuente de información, aunque se observa una gran variabilidad en su uso e intensidad entre las diferentes disciplinas. La investigación también confirma que las herramientas bibliométricas diseñadas para el análisis de la literatura

convencional son válidas en el análisis de la LG; no obstante, los parámetros todavía tienen que ser codificados para la evaluación del impacto de la LG sobre la literatura convencional en diferentes sectores. El trabajo de Cesare y Ruggieri (2006) compara tres parámetros (la frecuencia de uso de LG, la frecuencia de citación de LG y la intensidad de uso de LG) y se demuestra que solo el parámetro de la intensidad de uso ofrece una clara idea del impacto que puede tener la LG en un sector determinado y permite una comparación fiable entre los diferentes sectores. Utilizando el parámetro de frecuencia de uso, las autoras concluyen que el impacto de LG es más o menos igual en el campo de la Física y en el de Ciencias de la Información; del mismo modo, utilizando solo el parámetro de frecuencia de citación, no hay diferencia entre Ciencias de la Información y Demografía; por otro lado, usando el parámetro de intensidad de uso de LG, que tiene en cuenta tanto la frecuencia de uso como la frecuencia de citación, se obtienen diferencias significativas en ambos casos.

En cuanto al impacto de la literatura gris en la *Web*, Cesare y otros (2008) realizaron un análisis de citación de los diferentes tipos de LG utilizando *Google Scholar*. Los autores concluyen que el sistema de rastreo de las citas de *Google Scholar* da visibilidad a la literatura gris, tanto como documento citado como que cita, confirmando su valor en la comunicación académica. Además *Google Scholar* también permite reconstruir el ciclo de vida del documento al hacer posible que se encuentren las diferentes versiones a texto completo del documento, haciendo así evidente la transformación de la literatura gris en literatura convencional.

En la misma línea de análisis del impacto de la literatura gris, el estudio de Sibbald, MacGregor, Surmacz, y Wathen (2015) describe el proceso de desarrollo y aplicación de un “análisis modificado de citación” para la literatura gris. Este análisis se basa en los métodos que existen para examinar el impacto de un trabajo de investigación de dos formas: la primera, incluyendo deliberadamente literatura gris en el proceso de búsqueda de análisis de citas y, la segunda, mediante la inclusión de métodos cuantitativos y cualitativos de análisis para tener una mejor

comprensión de cómo se utilizó un trabajo de investigación. Al ampliar la búsqueda y profundizar en el nivel de análisis, se sugiere que este nuevo enfoque puede evaluar mejor el impacto de un trabajo, tanto si la investigación está dentro como fuera de los canales tradicionales revisadas por pares. Los autores comienzan con una revisión de la literatura gris y, a continuación describen los métodos actuales para analizar el impacto de un trabajo de investigación. Asimismo, los autores utilizan un ejemplo específico para describir el nuevo enfoque, que resalta su potencial para desarrollar el campo de análisis de citas e impacto, y discutir futuras mejoras y evaluación. Este estudio se suma al debate acerca de la inclusión y el análisis de los recursos de literatura gris en la búsqueda de diferentes tipos de "evidencias" para incluir tanto en contextos de investigación académica, de práctica y de políticas académicas. El estudio está dirigido a fomentar un enfoque más integral a la hora de evaluar el impacto de la literatura gris como nueva evidencia de la investigación.

Por otra parte, existen líneas de trabajo dirigidas a la acción y a la formulación de políticas que aumenten la parte de tesis doctorales electrónicas en libre acceso. En este sentido, el trabajo de Schöpfel y otros autores (2014a) argumenta que las e-tesis doctorales deberían estar accesibles libremente para servir mejor a los avances científicos, e informar a la comunidad científica en su conjunto sobre los resultados. Esto ya está apoyado técnicamente en las universidades por sus repositorios institucionales, por los archivos abiertos a nivel mundial, y por algunos editores. A través de los años, el porcentaje de las e-tesis doctorales que son de acceso abierto está aumentando en todo el mundo y en todas las disciplinas científicas. Abierto, debería ser por defecto y no una excepción. Sin embargo, la realidad es bien diferente. Incluso en los repositorios institucionales, creados para proporcionar acceso a la producción científica de una institución académica y como un sector central de la denominada vía verde para abrir el acceso, se echan en falta las partes más o menos importantes de la producción científica; debido a la falta de concienciación, al embargo, al depósito de metadatos sin texto completo, al contenido

confidencial, etc. Las tesis doctorales electrónicas, en particular, se difunden con diferentes tipos de estado, algunas son de libre disposición, otras están bajo embargo, bajo confidencialidad, bajo acceso restringido al campus (encriptado o no) o simplemente no están disponibles de ninguna forma. Contrariamente a la filosofía *Open Access*, muchos repositorios institucionales no son totalmente de acceso abierto (Prost & Schöpfel, 2014). Pero mientras que otros tipos de trabajos podrían estar disponibles a través de canales alternativos (revistas, monografías, etc.), las tesis doctorales electrónicas la mayoría de las veces no lo están.

El trabajo de Schöpfel y otros autores (2014a) resume la evidencia empírica de diferentes estudios y describe algunos de los principales resultados de una pequeña encuesta franco-alemana. Se propone un enfoque conceptual diseñado no solo para una mejor comprensión de las restricciones de acceso, sino también para la puesta en marcha de políticas que aumenten la parte de las tesis doctorales electrónicas en libre acceso. Al final, el documento presenta el esbozo de un proyecto de infraestructura europea llamada "Tesis y Disertaciones Electrónicas de Acceso Abierto" (ETD4OA) que se ha presentado en el programa marco *Horizon 2020*.

Capítulo I.7. Bibliometría para el acceso abierto. Métricas alternativas

Desde hace años está en curso un debate sobre la necesidad de revisar las métricas que evalúan la excelencia investigadora. Internet y otros desarrollos tecnológicos permiten la obtención de datos más ricos y la posibilidad de nuevos enfoques para evaluar la investigación.

El acceso abierto aboga por maximizar el impacto de la investigación mediante la mejora de la accesibilidad. En este sentido, los repositorios y las revistas de acceso abierto han venido utilizando nuevas herramientas y estrategias para mostrar cómo se puede beneficiar la ciencia de la diseminación libre en línea.

Lo más reciente en este debate de las métricas son las iniciativas basadas en las altmétricas o métricas alternativas, que vienen a enriquecer el paisaje con indicadores prometedores.

En este capítulo se ofrece una visión general sobre las métricas y el papel del acceso abierto en la promoción de un nuevo escenario para evaluar la ciencia, que servirá de marco para la parte de la investigación de esta tesis sobre la medición de la visibilidad, uso y citación de las e-tesis en acceso abierto.

I.7.1. Bibliometría. Definición y terminología

La bibliometría es el análisis cuantitativo de los patrones de publicación dentro de un determinado campo o área de la literatura (Brody, 2006).

Según algunos autores la bibliometría puede definirse sintéticamente como la disciplina responsable de la medición de la comunicación y, de forma más amplia, como la especialidad responsable de estudiar cuantitativamente la producción, distribución, difusión y el consumo de información contenida en cualquier tipo de documento (libro, revista, conferencias, tesis,

patentes, o sitio web) y cualquier campo intelectual, pero con especial atención a la información científica (Martin-Martin, Orduna-Malea, Ayllon, & Lopez-Cozar, 2016)

Los mismos autores afirman que es una disciplina con características peculiares, señalando en especial tres:

1. Se trata de una disciplina reciente, aunque enraizada a principios del siglo XX en el entorno de la biblioteca con la idea de medir la producción de conocimiento (estadísticas bibliográficas), y gestionar adecuadamente las colecciones de bibliotecas, no es hasta después de la Segunda Guerra Mundial que la bibliometría comienza realmente a establecer sus bases. Sus fundamentos epistemológicos aún están en ebullición.
2. Es una disciplina que se define mejor por sus métodos que por las áreas temáticas que cubre. Las denominadas "métricas" consideradas como el análisis de datos cuantitativos que aplican diversas técnicas estadísticas.
3. Tiene un fuerte carácter interdisciplinario que surge de la incorporación de métodos y técnicas desarrolladas en otros campos y por su aplicación al estudio de cualquier materia. Esto hace que la bibliometría sea una disciplina abierta, dispuesta para ser enriquecida por las ideas de la más diversa procedencia y para aceptar los científicos de los más diversos entornos disciplinarios. Esta es la razón por la que la bibliometría se asemeja a un cruce de caminos, un lugar donde se reúnen diferentes tradiciones científicas.

Para Hood & Wilson (2001) ha habido una notable confusión con la terminología de los tres términos métricos estrechamente relacionados: bibliometría, cienciometría, e informetría. Los autores realizan una extensa discusión sobre estos términos, considerando la webometría (estudios de la Web) como un subcampo de la cibermetría (estudios de todas las aplicaciones de Internet), aunque a menudo estos dos términos se consideren sinónimos.

Bar-Ilan (2008) define la informetría como el estudio cuantitativo de los aspectos de la información en cualquier forma, no solo registros o bibliografías, y en cualquier grupo social, no solo científico. Egghe (2005) utiliza el término informetría como el término amplio que comprende todos los estudios métricos relacionados con la ciencia de la información, incluyendo la bibliometría (bibliografías, bibliotecas, etc.), la cienciometría (ciencia políticas, análisis de citación, evaluación de la investigación, etc.), la webometría (métrica de la web, Internet u otras redes sociales como las redes de colaboración o citación, etc.).

Para Martín-Martín, Orduna-Malea, Ayllón y López-Cozar (2016) el hecho de que la bibliometría sea una disciplina reciente y de carácter interdisciplinario e instrumental hace que se la conozca por muchos nombres diferentes. Sin embargo, esto no significa que el objeto de estudio o las fronteras de la disciplina no estén claramente definidos. Más bien, es un signo de la coexistencia de diferentes tradiciones que han dado forma al desarrollo de la disciplina.

La bibliometría es el nombre original y el más extendido. Es fruto de la tradición bibliográfica representada por Paul Otlet (1934) con su propuesta de "bibliometrie", una ciencia para medir todos los aspectos de los libros y otros documentos, y de la tradición de la biblioteca desde la antigüedad sobre la medición del crecimiento del conocimiento y el uso de sus fondos.

La cienciometría está orientada hacia el análisis cuantitativo de la literatura científica y técnica. Proviene de la tradición de la ciencia de la ciencia (espacio de confluencia de Sociología, Historia y Filosofía de la ciencia), a la que la política científica también está vinculado. La creación de los índices de citas (bases de datos dedicadas a la recogida de la producción científica) fue crucial para esta orientación cienciométrica.

La informetría se centra en el descubrimiento de los modelos matemáticos que explican las propiedades de información. Está conectada con la ciencia moderna de la información. Es una denominación tan próxima a cienciometría que a veces es difícil encontrar diferencias entre ellas.

Webométrica y alométrica son las denominaciones más recientes. Ambas comenzaron a ganar impulso a medida que comenzó a extenderse el uso de las nuevas tecnologías de la información y de la comunicación. Las dos se desarrollan en la tradición de la Biblioteca y las Ciencias de la Información, una disciplina cada vez más dedicada a la informática y a la informatización de ella misma. Vienen también para resaltar el aspecto tecnológico tradicional del que las diferentes especialidades métricas han disfrutado desde su inicio.

Un análisis llevado a cabo por Martín-Martín y otros autores sobre los términos utilizados en los títulos de los documentos en el área de la Biblioteconomía y las Ciencias de la Información, publicados entre 1969 y 2015 e indexados en *Google Scholar*, muestra un claro predominio de la expresión "bibliometría" seguida de "cienciometría". Sin embargo, en los últimos tres años, el término "altétrica" se está utilizando cada vez más, como resultado de la novedad de las nuevas tecnologías de la comunicación de los medios sociales.

Un resultado similar se obtiene cuando se analizan las palabras clave utilizadas por los 814 científicos con un perfil público en Google Académico, especializados en bibliometría o que trabajan esporádicamente en este campo. La prevalencia de la cienciometría y bibliometría es clara, aunque el peso de esta última sería mayor si los términos hubiesen sido correctamente normalizados (Martín-Martín et al., 2016). En la Figura 25 se muestra la nube de etiquetas con los términos usados por los investigadores con un perfil público en *Google Scholar Citations*; la frecuencia de uso se indica por el tamaño en la muestra.



Figura 25. Nube de etiquetas con términos usados por los investigadores con un perfil público en *Google Scholar Citations*. Fuente: Martín-Martín, Orduna-Malea, Ayllón, & Delgado López-Cózar (2016)

1.7.2. La comunicación científica en línea y la bibliometría tradicional

Acuñada por Eugene Garfield en 1955 *Journal Impact Factor* (JIF) ha sido una medida predominante utilizada para evaluar la literatura científica e identificar el prestigio, la promoción y la relevancia de los investigadores. Originalmente creado para ayudar a seleccionar las revistas a indizar en los *Science Citation Index*, este indicador sigue ejerciendo una gran influencia en los planes estratégicos científicos, en las decisiones de financiación y en la comunicación académica.

Aunque su creador Garfield esté en contra de su uso para evaluar a personas e instituciones (Garfield, 2006), aunque se hayan propuesto decenas de fórmulas para retocarlo y de indicadores para sustituirlo, el hecho es que el JIF ha sido el auténtico patrón de oro de la bibliometría (Orduna-Malea, Martín-Martín, & Delgado-López-Cózar, 2016).

El factor de impacto de las revistas fue ideado como una medida cuantitativa de la importancia de una revista, normalizado por el tamaño de la revista (por la Ley de Bradford las revistas varían

en gran medida en el número de documentos que contienen). El JIF ha cobrado importancia tanto por la presión de las cancelaciones de las bibliotecas, que pueden utilizar el JIF para determinar qué revistas son menos "importantes" para cancelar, sino también porque se ha utilizado en la evaluación de los autores (Brody, 2006).

Sin embargo, el mal uso del JIF y la necesidad de establecer una mejor correlación entre este indicador y otras métricas de rendimiento de la investigación es prácticamente una novedad, tanto más cuanto que internet y otros desarrollos tecnológicos han desatado su potencial. Abundan las críticas al monopolio de *Thomson Reuters* sobre el factor de impacto y del mismo modo se han puesto en tela de juicio sus procesos secretos y no científicos para clasificar las revistas especializadas en el índice (Bernal, 2013).

Han surgido nuevos indicadores como el índice h (2005)³⁶ que se centra en la productividad y el impacto de las obras de los autores con mayor precisión (Hirsch, 2005), mientras que el *ranking* de la revista, como Eigenfactor (2007)³⁷, pondera la importancia de las citas en función de la relevancia de las revistas para la comunidad científica (Bergstrom, West, & Wiseman, 2008); y además también ha sido impulsada por Elsevier Scimago (2007) y el SNIP (*Source Normalized Impact per Paper*, 2010), indicadores que tienen en cuenta la reputación en línea de las puntuaciones de estos indicadores. El indicador *SCImago Journal Rank* (SJR) está diseñado para ser usado por complejas y heterogéneas redes de trabajo como *Scopus* (González-Pereira, Guerrero-Bote, & Moya-Anegón, 2010). SNIP mide el impacto de la citación contextual de una revista, teniendo en cuenta principalmente la frecuencia con la que los autores citan otros artículos en las listas de referencia y la rapidez de madurez de la cita (Moed, 2010).

No obstante, el número de citación académica y el prestigio de las revistas entre pares continúan sustentando estos modelos (Bernal, 2013).

³⁶ Índice h, https://es.wikipedia.org/wiki/%C3%8Dndice_h

³⁷ Eigenfactor, http://www.eigenfactor.org/projects/researchArea_1.php

Bernal (2013) comenta que la llegada de la tecnología de la información electrónica, particularmente la de Internet, abre la vía a nuevas revistas, nuevos datos y nuevos enfoques para la evaluación de la investigación. Desde 1990 la relación entre JIF y la citación de los artículos ha ido decreciendo gradualmente a través de las disciplinas. El acceso abierto ha hecho que el debate de la revalorización de los indicadores de impacto científico sea más evidente ya que en general la accesibilidad y el fácil intercambio de información académica han allanado el camino para nuevas y más profundas consideraciones a nivel científico.

La necesidad de nuevos sistemas de evaluación fue llevada a la Conferencia BOAI en 2012 y se convirtió muy pronto en un *hot topic* entre los financiadores, los editores, los promotores del acceso abierto, los administradores de repositorios y los bibliómetros, así como en el sector de la investigación y la innovación, y la comunidad del *software*.

I.7.3. Nuevas herramientas y métricas de la ciencia en un escenario de acceso abierto

Por desgracia todavía no hay estándares para los indicadores de uso por lo que muchos repositorios publican extensos informes sobre las visitas y las descargas, datos que no pueden ser comparados y combinados como conceptos medidos ya que no son lo mismo (Aguillo et al., 2010).

Los esfuerzos actuales por desarrollar medidas de citación y uso basados en indicadores se están desarrollando (Citebase: Brody, 2003; CiteSeer: Lawrence, Giles & Bollacker, 1999; RepEc Ideas: Zimmermann, 2007), pero al mismo tiempo se necesita una primera generación de métricas de los repositorios de acceso abierto por dos razones fundamentalmente, por un lado para reforzar las iniciativas de acceso abierto, la promoción de depósito de documentos por los investigadores y por otro lado para proporcionar datos empíricos acerca de este nuevo sistema

de comunicación académica para su posterior análisis y comparativa. En este sentido el *Ranking Web de Repositorios*³⁸ ha desarrollado una serie de indicadores para medir su actividad y visibilidad y presentar los resultados obtenidos sobre actuales y futuros desarrollos.

Al comienzo de la era digital hubo una ruptura gradual de la correlación entre las revistas altamente citadas con los artículos más citados. Sin embargo, a excepción de las áreas específicas de investigación, como la física, el salto del papel a la publicación en línea no causó una necesidad inmediata de adoptar nuevos tópicos como los de compartir o evaluar la ciencia.

El uso del acceso abierto y de herramientas de la web 2.0 en la última década ha contribuido a producir transformaciones profundas en las bases del modelo. En la actualidad, los investigadores están involucrados activamente en plataformas como *Mendeley*³⁹, *ResearchGate*⁴⁰ y *Academia*⁴¹ con el fin de hacer que sus trabajos sean más fácilmente localizados, puedan ser comentados y establecer conexiones más rápidas con sus colegas. Los investigadores también autoarchivan sus artículos postprint y preprint en repositorios institucionales y temáticos, publican en revistas de acceso abierto y cada vez más frecuentemente hacen públicos los datos en bruto de su investigación científica.

La supresión de las barreras económicas y técnicas para buscar, recuperar y acceder al texto completo de la literatura científica fomenta el aumento de un mayor número de lectores. Y además como valor añadido, el acceso abierto también puede aumentar el impacto científico.

Fue por primera vez en 2001 que se informó que los *proceedings* científicos disponibles libremente *online* acumulaban tres veces más el número de citas recibidas por los artículos impresos (Davis et al., 2008). Esta "ventaja de citación" se ha encontrado en muchas disciplinas y en un estudio de 2010 se resumen las principales conclusiones y metodologías de los

³⁸ Ranking Web de Repositorios, <http://repositories.webometrics.info>

³⁹ Mendeley, <https://www.mendeley.com/>

⁴⁰ ResearchGate, <https://www.researchgate.net/>

⁴¹ Academia, <https://www.academia.edu/>

principales estudios (Alma Swan, 2010). El proyecto OpCit⁴² rastreaba exhaustivamente la literatura relacionada con el efecto del acceso abierto y las descargas sobre el impacto de citación; este proyecto ya no está disponible tras el cierre de Connotea, en marzo de 2013⁴³.

Por término medio, los artículos disponibles libremente tienden a ser más citados y a veces esta tendencia ha sido relacionada con el llamado “postulado de autoselección”, lo que significa que los autores autoarchivan sus trabajos que han sido publicados en revistas con alta tasa de citación (Antelman, 2004). Sin embargo, estudios recientes señalan que la ventaja de citación del acceso abierto se produce independientemente de otras variables como la edad del artículo, el factor de impacto de la revista, número de coautores, referencias o páginas, campo de investigación, tipo del artículo o el país de publicación (Gargouri et al., 2010). Ese tipo de ventaja alcanza el punto más alto para los artículos más citados. Además el acceso abierto está probando modelos alternativos para los *rankings* de revistas científicas. Por ejemplo, teniendo en cuenta la apertura relativa de las políticas de los editores en relación con el autoarchivo, un equipo de trabajo del MIT está trabajando en el llamado “cociente de accesibilidad”. Esta medida se calcula combinando el *Ted Bergstrom's Relative Price Index* que evalúa la asequibilidad y la calidad, con los datos de SHERPA/RoMEO que evalúa el derecho a compartir la versión *peer-reviewed* de un artículo (Willmott, Dunn, & Duranceau, 2012). Otra herramienta, *Google Scholar Metrics*, intenta clasificar las revistas mediante la recopilación de citas en un amplio conjunto de fuentes, incluyendo los repositorios y plataformas no tradicionales⁴⁴.

Tanto la vía dorada como la vía verde del acceso abierto han desarrollado nuevas estrategias y herramientas para medir la atención que suscitan los resultados de investigación. Un buen artículo puede llevarle entre 2 y 5 años iniciar la recopilación de citas académicas. Sin embargo mientras tanto, otros instrumentos pueden cuantificar su impacto y ayudar a hacer estimaciones

⁴² OpCit Project, <http://opcit.eprints.org/oacitation-biblio.html>

⁴³ Connotea, <http://www.connotea.org/>

⁴⁴ Google Scholar, http://scholar.google.com/citations?view_op=top_venues

sobre el crecimiento potencial de las citas en el futuro. Por tanto la medición del impacto a corto plazo a través de conteos de visitas HTML, descargas, compartidos sociales y marcados en gestores de referencias científicos podrían ser indicativo del impacto potencial a medio y largo plazo.

La utilización de repositorios ha facilitado la implementación del uso de módulos de estadísticas como los desarrollados por Eprints y DSpace. Algunos aportan gran nivel de detalle, ya que proporcionan información sobre la distribución demográfica relacionado con el uso, las principales palabras clave de búsqueda, los datos agregados sobre el comportamiento del usuario final, la relación del tráfico generado dentro de la institución y en otros lugares, etcétera.

Debido a su popularidad y la facilidad de uso, la mayoría de los repositorios suministran datos relativos a visitas y descargas gratuitas de texto completo. Los investigadores hacen uso de estos datos para obtener una mejor visión del alcance de sus obras que están en acceso abierto, mientras que los administradores de repositorios se basan en los datos para prever las estrategias de desarrollo de contenidos, impulsar el uso, promover colecciones similares a los recursos más descargados, y presentar pruebas de un mayor impacto de los contenidos de acceso abierto a los equipos de gobierno institucionales. Sin embargo, existe una cuestión pendiente en el ámbito de las estadísticas de uso referida a la falta de una norma internacional que haga posible la recopilación de datos de uso a través de plataformas editoriales, *middleware*, repositorios de acceso abierto y otros sitios de manera consensuada. Un ejemplo positivo es el proyecto IRUS⁴⁵ en el Reino Unido, que provee estadísticas de uso de acuerdo a COUNTER para todo el contenido descargado de los repositorios institucionales participantes del Reino Unido.

⁴⁵ IRUS-UK, <http://www.irus.mimas.ac.uk/>

En menor medida, los repositorios configuran herramientas para rastrear el número de citas en un entorno abierto. En este sentido, los repositorios temáticos llevan la delantera y por ejemplo CiteSeer, RePEC, arXiv y SSNR (adquirido recientemente por Elsevier) mantienen métricas para los trabajos alojados en sus plataformas y citados por otras disponibles también en acceso abierto. Los repositorios institucionales han seguido su ejemplo, y la mayoría de las veces se han añadido APIs que muestran las citas académicas como las obtenidas por WOS, Scopus y *Google Scholar*. *Open Citations Project* exploró cómo publicar la información bibliográfica y la cita de la literatura biomédica en RDF con el fin de facilitar los enlaces de las citas. Su actualización está trabajando en un prototipo para un uso más amplio⁴⁶

Al centrarse en las métricas de un solo artículo, el acceso abierto ha puesto el énfasis en la importancia del impacto de los resultados de la investigación en sí, no en el prestigio de las revistas donde se publican.

Un modelo que ha sido ampliamente adoptado por las iniciativas de acceso abierto y los editores basados en la suscripción es el desarrollado por PLoS, <https://www.plos.org/>. Dado que el número de citación evalúa el impacto académico a medio y largo plazo, el efecto inmediato de las obras a través de múltiples lugares y foros suele descuidarse bastante. Para llenar el vacío, las métricas de nivel artículo de PLoS incluyen, además de los puntos de referencia tradicionales, como las citas académicas, otros datos como las citas no académicas (blogs científicos como ScienceSeeker o Wikipedia), estadísticas de uso y métricas alternativas (compartición social y marcadores científicos).

Plum Analytics⁴⁷ funciona de forma similar a PLoS. Clasifica el impacto de los artefactos de investigación (desde artículos a vídeos) en cinco categorías diferentes: uso, capturas, menciones, medios sociales y citaciones. El repositorio Gredos en su nueva versión incorporará

⁴⁶ OC OpenCitations project, <http://opencitations.net/>

⁴⁷ Plum Analytics, <http://plumanalytics.com/>

esta herramienta así como otras citadas anteriormente en este capítulo (APIs de WOS, Scopus, Google Scholar) con el fin de incorporar métricas alternativas que informen del uso de todas los trabajos contenidos en el repositorio. Aunque estas herramientas son muy recientes existe un gran entusiasmo con su utilización, no solo por los repositorios institucionales de acceso abierto sino también por los editores (Cambridge University Press, BioMed, Nature, Taylor&Francis), las bases de datos bibliográficas (Scopus, WOS).

A través de algunos estudios sobre el impacto de la investigación de acceso abierto y la medición del número de citas se han identificado las tendencias más significativas como:

1. La ventaja de citación del acceso abierto se sitúa entre el 25% and 250% dependiendo de la disciplina. Esta ventaja está impulsada por el valor intrínseco de las obras y no por el mero hecho de ser de libre acceso (Wagner, 2010).
2. El acceso abierto, ya sea a través de autoarchivo voluntario o por mandato, tiene un efecto positivo sobre el impacto científico de nuevos artículos (Gargouri et al., 2010).
3. Puede haber una correlación entre la difusión de la investigación a través de la web social y las descargas de los repositorios (Terras, 2012).
4. Las descargas de artículos HTML/PDF junto con las comparticiones en medios sociales son mejor indicador del interés del público que las estadísticas de uso absoluto (Fenner & Lin, 2014; J. Lin & Fenner, 2013).
5. Como indicadores complementarios, las métricas académicas pueden mostrar un valor añadido para el contenido de libre acceso, y este hecho podría actuar como un nuevo incentivo para que los autores aumenten sus tasas de autoarchivo en el repositorio (Konkie & Scherer, 2013).
6. Para el contenido en acceso abierto, a menudo parece haber una correlación entre páginas vistas y descargas de texto completo con citas académicas y el impacto social

en los espacios sociales (Tim Brody, Harnad, & Carr, 2006; Jason Priem, Piwowar, & Hemminger, 2012).

7. La citación académica está a menudo estrechamente relacionada con el uso de marcadores sociales (Haustein & Siebenlist, 2011).
8. Se está incrementando el uso de marcadores sociales y gestores de referencias (Mendeley, Research Gate, CiteuLike) y son la mejor medida del interés académico por los medios sociales (Twitter, Facebook) donde el interés público es más fuerte (Fenner & Lin, 2014; J. Lin & Fenner, 2013).
9. Los *posts* de autores en blogs son especialmente significativos para la comunidad científica en la medida en que proporcionan comentarios y crítica de artículos publicados recientemente, enriqueciendo así las obras originales (Bernal, 2013).

I.7.3.1. Las métricas alternativas o alométricas

Las métricas alternativas o métricas de citación alternativa ofrecen nuevas formas de rastrear la influencia de los autores a través de una amplia gama de medios y plataformas (Galloway & Pease, 2013). Jason Priem uno de los líderes en esta área define las alométricas como el estudio de las medidas de impacto basado en la actividad de las herramientas y los entornos *online* (Priem, Groth, & Taraborelli, 2012). Las alométricas van más allá del artículo de investigación para rastrear la influencia de otras formas científicas definidas por la *National Science Foundation* (NSF)⁴⁸ como productos citables y accesibles no limitados a publicaciones, conjuntos de datos, *software*, patentes y derechos de autor.

Los métodos tradicionales del impacto científico se basan en el recuento de citas y han sido la norma por la cual se mide el impacto de la investigación. Estos métodos siguen siendo

⁴⁸ NSF, <http://www.nsf.gov/>

importantes, incluso para fines de promoción y de prestigio, pero no proporcionan una visión completa.

La recopilación de los datos es lenta y, a menudo estos no son públicos y se encuentran dentro de sistemas propietarios. Tampoco se tienen en cuenta la diversidad de mecanismos de publicación disponibles, ni los tipos y formas bajo los que se publica la investigación, como por ejemplo la literatura gris (informes técnicos, documentos de trabajo e informes técnicos), la cual tampoco está bien representada en los sistemas bibliométricos tradicionales. Dado que la literatura gris se publica cada vez más con más frecuencia directamente en la web, la producción académica que contiene es ahora más fácil de encontrar y citar, por lo que es una buena candidata para su inclusión en el cuerpo del trabajo científico utilizando métricas alternativas.

Además, las altmétricas son utilizadas por los científicos comprometidos como un tipo de servicio de asesoramiento a los lectores, proporcionando apoyo a la investigación y sugerencias. El uso de herramientas de los medios sociales como las que se describen a continuación, posibilita seguir a otros expertos del mismo campo científico, unirse a grupos de interés y compartir tanto las referencias como los resultados de la investigación. En un sentido, algunas de estas interacciones se desarrollan casi en tiempo real de forma que cualquier persona pueda tomar parte en las conversaciones permitiendo así avanzar en el conocimiento.

I.7.3.2. Las herramientas de las altmétricas

Los métodos de diseminación digital de la investigación y de la comunicación científica continúan evolucionando. Las nuevas métricas han tomado los canales por los que los artículos se difunden a través de un ecosistema científico en expansión, llegando más allá del alcance de los rastreadores y filtros tradicionales. Mediante la monitorización y la captura de la huella de la investigación desde el momento de la publicación y conforme circula por toda la comunidad, las

métricas alternativas o alométricas van midiendo también el impacto agregado de la propia investigación.

Las métricas alternativas son el estudio y uso de las medidas de impacto basadas en la actividad de los entornos web. La investigación con alométricas busca construir y realizar un seguimiento total en diversas audiencias (tanto público en general como académico) y monitorear los diversos tipos de compromiso con el mundo académico, incluyendo la audiencia, la discusión, etc.

La aparición de estas nuevas métricas se ha enmarcado en la denominación *altmetrics* o “métricas de los medios sociales” introducida en 2010 por Priem y otros autores (Priem, Taraborelli, Groth, & Neylon, 2010) como una vía alternativa para medir el impacto de la investigación en la web social a través de diferentes herramientas (Priem, Groth, et al., 2012; Priem, Piwowar, et al., 2012). Más específicamente, las alométricas recogen las menciones de los resultados científicos en los medios sociales, en nuevos medios y en gestores de referencias. Este desarrollo del concepto de alométricas ha sido acompañado por un aumento en la diversidad de herramientas que permiten rastrear en tiempo real el impacto de los resultados científicos explorando los “compartidos” (*shares*), los “me gusta” (*likes*), los “comentarios” (*comments*), las revisiones (*reviews*), las “discusiones” (*discussions*), los “favoritos” (*bookmarks*), los “guardados” (*saves*), los *tweets* y las menciones de las publicaciones científicas y fuentes en los medios sociales (Wouters & Costas, 2012; Zahedi et al., 2014).

Galloway y Pease (2013) consideran que *Google Scholar Citations* es una de las herramientas más maduras y prometedoras de las alométricas, destinada a rastrear el número de lectores y la influencia. Los autores la describen como un servicio que permite a los autores rastrear sus publicaciones e influencia usando *Google Scholar Metrics*. Se trata de una herramienta libre, extremadamente útil, de uso amigable y bien considerada.

Entre las herramientas y servicios más destacables se encuentran (Figura 26):

F1000	http://f1000.com/
Impactstory	https://impactstory.org/
Almetric.com	https://www.altmetric.com/
Plum analytics	http://plumanalytics.com/
ReaderMeter	http://readermeter.org/
PeerEvaluation	http://www.peerevaluation.org/
ScoreCard	http://researchscorecard.com/
PLOS Article-Level-Metrics (ALM)	http://article-level-metrics.plos.org/
Citeulike	http://www.citeulike.org/home
Mendeley	www.mendeley.com



Figura 26. Herramientas de las altmétricas. Fuente: Elaboración propia.



MÓDULO II. SISTEMA DE ANÁLISIS Y CASO DE ESTUDIO

Módulo II. Sistema de análisis y caso de estudio

En el módulo 2 “Sistema de análisis y caso de estudio” se recogen los componentes, las funciones y características del protocolo y del sistema utilizado para la obtención de datos e informaciones relativas a la investigación llevada a cabo, así como la presentación, descripción y características del caso estudiado: las e-tesis en acceso abierto a través del repositorio Gredos.

El módulo se compone de dos capítulos. En el primero se presenta y se describe el caso de estudio tratado en esta investigación, es decir la comunidad “Tesis doctorales” del repositorio Gredos. En el segundo capítulo, titulado “Material y métodos”, se presenta el material y métodos utilizados, refiriéndose a la muestra, la selección de material, la definición de las variables, el análisis y los métodos estadísticos empleados. El módulo tiene la siguiente estructura:

Módulo II. Sistema de análisis y caso de estudio

Capítulo II.1. Caso de estudio: Tesis doctorales del repositorio Gredos

- II.1.1. Colección de Tesis doctorales en el repositorio Gredos
- II.1.2. Normativa sobre el depósito de tesis en Gredos
- II.1.3. Depósito de las tesis doctorales en Gredos
- II.1.4. Disseminación de las tesis doctorales de Gredos: la interoperabilidad

Capítulo II.2. Material y métodos

- II.2.1. Muestra
- II.2.2. Selección del material
- II.2.3. Definición de las variables
- II.2.4. Análisis y métodos estadísticos

Capítulo II.1. Caso de estudio: tesis doctorales de Gredos

Los estudios de caso son informes de materiales de caso obtenidos mientras se trabaja con un elemento individual, un grupo, una comunidad, o una organización. Los estudios de caso ilustran un problema, indican una forma de solucionarlo, y aportan luz sobre la investigación, sobre las aplicaciones prácticas, o sobre las cuestiones teóricas. Al escribir estudios de caso, los autores tienen en cuenta el equilibrio entre el suministro de material ilustrativo importante y el uso de material confidencial del caso de manera responsable (The Writing Lab, The OWL at Purdue, & Purdue University, 2016).

En este capítulo se realiza una descripción de la colección de Tesis doctorales de Gredos, aportando datos numéricos, la evolución del depósito de las tesis doctorales y la normativa a la que están sujetas. Se realiza una breve exposición sobre los procesos del sistema de carga y la descripción de las tesis en Gredos. Por último se hace referencia a la disseminación y difusión de esta colección gracias a la interoperabilidad del repositorio.

Como ya se ha expuesto anteriormente, la elección de las tesis doctorales del repositorio Gredos como caso de estudio de esta investigación, está motivado porque las tesis doctorales constituyen la única investigación sobre la que la Universidad de Salamanca estableció un mandato a partir del curso 2008-2009 para su depósito y difusión en abierto a través del repositorio institucional. Por ello, de la colección de tesis en abierto disponibles Gredos solo son objeto de estudio las tesis doctorales leídas entre los años 2008 a 2011, período en que está en vigor esta normativa.

II.1.1. Colección de Tesis doctorales en el repositorio Gredos

La comunidad Tesis doctorales de la Universidad de Salamanca en Gredos se estructura en cinco colecciones, de acuerdo a las áreas de conocimiento a las que pertenezcan los departamentos, centros o institutos de investigación donde se haya realizado la investigación.

Las colecciones son las siguientes:

- TD. Ciencias biosanitarias
- TD. Ciencias experimentales
- TD. Ciencias sociales
- TD. Enseñanzas técnicas
- TD. Humanidades

El número de tesis depositadas en acceso abierto a finales de 2011 superaba las 800 tesis, y cinco años después, en 2016, el número asciende a 2.100, como se puede comprobar en la imagen (Figura 27).

Repositorio Documental
GREDOs

Repositorio Documental de la Universidad de Salamanca >
Repositorio Científico >
Tesis >

Buscar en Gredos
Todo Gredos

→ Búsqueda avanzada
→ Áreas temáticas
→ Página de inicio

Comunidad Tesis doctorales : [2100]
Página de inicio de la comunidad

En:
Tesis doctorales

Buscar por:
Ir

Visualizar los registros ordenados por:
Fecha Publicación Autor Título Materia

Esta comunidad contiene las tesis doctorales de las distintas áreas de conocimiento de la Universidad de Salamanca. Además contiene también una colección de tesis no pertenecientes a la USAL, aunque guardan relación con esta institución académica.

Colecciones en esta comunidad

- TD. Ciencias biosanitarias [635]
- TD. Ciencias experimentales [248]
- TD. Ciencias sociales [737]
- TD. Enseñanzas técnicas [80]
- TD. Humanidades [398]
- Tesis doctorales no USAL [2]

Envíos recientes

- Estudio de la memoria de trabajo en sujetos adultos portugueses
- Las isoformas de endoglin en la regulación de la angiogénesis
- Supervisão e aquisição da Competência Profissional: Um modelo de e-tutoring na formação de Educadores de Infância
- Optimización de un nuevo proceso electroquímico para descontaminación de residuos finos de escobrecera de mina
- A história da educação física em Portugal: da implantação da República (1910) à aprovação da Lei de Bases do Sistema Educativo (1986)

Recursos RSS

RSS 1.0 RSS 2.0 RSS

Figura 27. Comunidad Tesis doctorales del repositorio Gredos y sus colecciones. Fuente: repositorio Gredos (2016)

En el gráfico (Figura 28) se observa la relación del tamaño entre la colección de tesis doctorales con el resto de los tipos de documentos alojados en la comunidad Repositorio Científico.

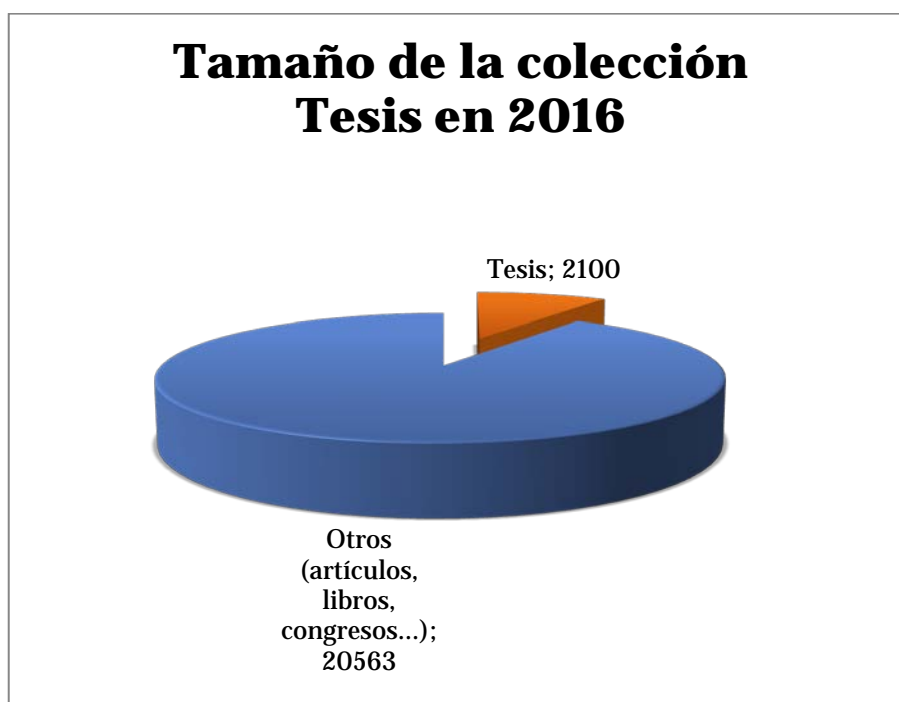


Figura 28. Tamaño de la colección de Tesis doctorales del repositorio Gredos (2016). Fuente: Elaboración propia

II.1.2. Normativa sobre el depósito de tesis en Gredos

La Universidad de Salamanca firmó su compromiso con el acceso abierto a través de la Declaración de Berlín en febrero de 2009. Este compromiso adquirido por la Universidad se vio plasmado en varios mandatos aprobados en Consejo de Gobierno de la Universidad de Salamanca de 29 de febrero de 2008 y modificada por los Consejos de Gobierno de 29 de septiembre de 2008 y 27 de febrero de 2009 (sección tercera, artículo 23), y que hacían referencia a la obligación de depositar la copia digital de las Tesis en el Repositorio Institucional. Por ello, las tesis doctorales que se presentaban a partir de la fecha de aprobación del mandato pasaban automáticamente al repositorio institucional sin requerir de la firma o aprobación de los autores. El repositorio institucional garantiza los derechos de autor, de acuerdo a la ley de Propiedad Intelectual, y toda la documentación contenida en Gredos tiene una licencia *Creative*

Commons 3.0 España (Ferrerías-Fernández & Merlo-Vega, 2010). Se trata del primer y único mandato institucional de acceso abierto de la Universidad de Salamanca.

Por otro lado, el día 10 de febrero de 2011 se aprobó la normativa sobre el depósito de tesis en acceso abierto expuesta en el Real Decreto 99/2011 que regula los estudios oficiales de doctorado y que establece en su artículo 14.5 el “mandato nacional” de depositar el archivo electrónico abierto de las tesis doctorales en el repositorio institucional correspondiente (Ministerio de Educación, 2011).

La entrada en vigor del Real Decreto 99/2011 viene a inhabilitar el mandato institucional de 2009, aunque en el artículo 14 se establezca el depósito de las tesis en acceso abierto en repositorios institucionales y esto constituya un mandato nacional de acceso abierto para las tesis leídas en las universidades españolas. Paradójicamente, en el caso de la Universidad de Salamanca no se ha reforzado el mandato que ya existía, sino que se ha visto limitado y debilitado, ya que si bien, en el punto 5 del citado artículo 14, se indica que *“una vez aprobada la tesis doctoral, la universidad se ocupará de su archivo en formato electrónico abierto en un repositorio institucional y remitirá, en formato electrónico, un ejemplar de la misma así como toda la información complementaria que fuera necesaria al Ministerio de Educación a los efectos oportunos”*, sin embargo, en el punto 6 del mismo artículo 14, se indica que *“en circunstancias excepcionales determinadas por la comisión académica del programa, como pueden ser, entre otras, la participación de empresas en el programa o Escuela, la existencia de convenios de confidencialidad con empresas o la posibilidad de generación de patentes que recaigan sobre el contenido de la tesis, las universidades habilitarán procedimientos para desarrollar los apartados 4 y 5 anteriores que aseguren la no publicidad de estos aspectos”*. Y aunque no son muchos los autores que se acogen a dicha excepción para no depositar su tesis en abierto, no deja de ser un retroceso en los logros anteriores alcanzados con respecto al fomento del conocimiento abierto y al compromiso adquirido con la firma de la Declaración de Berlín.

II.1.3. Depósito de las tesis doctorales en Gredos

El proceso seguido desde que se deposita la tesis hasta que se carga en el repositorio es el siguiente:

1. El doctorando hace el depósito de su tesis en la Escuela de Doctorado de la Universidad de Salamanca.

2. A partir de la entrada en vigor del Real Decreto 99/2011, el doctorando debe cumplimentar el impreso “Archivo de la tesis doctoral en formato electrónico abierto en el repositorio institucional de la Universidad de Salamanca”, donde indicará si autoriza a depositar en el repositorio una copia íntegra de su tesis, o si solicita a la Comisión Académica del Programa de Doctorado que “se proceda a garantizar la no publicidad de aquellos aspectos de la tesis doctoral sobre los que existen intereses editoriales, comerciales o industriales legítimos que pueden verse perjudicados con la publicidad de una copia íntegra de su tesis doctoral”; en este último caso deberá aportar un informe en el que se indiquen los intereses en conflicto y una versión fidedigna de la tesis doctoral, que coincida esencialmente con el texto completo, en la que no aparezcan los aspectos objeto de intereses en conflicto, para que sobre ella se realice el citado archivo de la tesis doctoral.

3. Una vez defendida y aprobada, la tesis doctoral en formato electrónico (bien la copia íntegra o bien la versión reducida) es enviada al Repositorio Gredos, cuyo personal se encarga de describir y subir la tesis al repositorio dentro de la comunidad “Tesis Doctorales”, en la colección correspondiente al área de conocimiento. Si se trata de una versión reducida, esta se depositará en la comunidad “Extractos y versiones parciales de tesis”.

En esta tesis doctoral no se recoge el caso de las versiones parciales de tesis, ya que es una situación que se da a partir del curso 2011-2012, con la entrada en vigor del Real Decreto

99/2011, y como se verá más adelante la muestra analizada recoge solo las tesis doctorales presentadas durante el período de mandato institucional, cursos 2008-2009, 2010-2011.

Desde 2008 hasta el año 2011, se depositaba en el repositorio la copia íntegra de todas las tesis doctorales presentadas en la Universidad de Salamanca. Durante ese período, si el autor tenía algún impedimento para difundir en acceso abierto su tesis doctoral debía remitir una solicitud al Director del Servicio de Bibliotecas exponiendo los motivos por los que solicitaba la retirada de la tesis del repositorio. En el repositorio se disponía de una base de datos donde se registraban los datos de las tesis doctorales retiradas, el motivo y la fecha de su retirada del repositorio.

De acuerdo a la información extraída de esta base de datos, se comprueba que los autores, en muy pocas ocasiones, se han mostrado reticentes a depositar sus tesis doctorales en el repositorio institucional Gredos. La mayoría de los autores (95%) ve en el repositorio una oportunidad de difundir su obra para asegurar los derechos de autor frente al posible plagio de las obras inéditas y desconocidas en el ámbito científico, además de asegurar el aumento de la visibilidad y por tanto la citación y el impacto de su trabajo. Solo un 5% de los autores han estado disconformes con la difusión de su trabajo en acceso abierto en el repositorio alegando fundamentalmente tres motivos:

1. No asignar un ISBN a su publicación que le sirva para concurso de méritos.
2. La existencia de compromisos con un editor para publicar la tesis doctoral, y que este retenga los derechos de explotación, no permitiéndole el archivo en repositorios institucionales de acceso abierto.
3. La tesis contiene información sensible que no puede ser difundida.

En cuanto a la asignación de números ISBN por entidades o personas que no son editores, la Agencia Española del ISBN regula esta circunstancia desde agosto de 2010. La Agencia además pone énfasis en explicar qué no es un ISBN:

- El ISBN no es obligatorio. El Real Decreto 2063/2008 del Ministerio de Cultura sobre el ISBN, publicado en el BOE número 10, del 12 de enero de 2009, en su “Disposición derogatoria única a)”, deroga la obligatoriedad de consignar el ISBN en las publicaciones, fijada por Decreto 2984/1972.

- La Agencia del ISBN no tiene como misión valorar la calidad de las publicaciones para las que se solicita un ISBN. La Agencia no tiene acceso a las publicaciones, sino a los datos que sobre ellas les envían los editores.

- Los derechos de propiedad intelectual no quedan protegidos por la obtención de un ISBN sino por la propia publicación de la obra o por su inscripción en el Registro de la Propiedad Intelectual.

El depósito de tesis doctorales de la Universidad de Salamanca en el repositorio Gredos es una forma de difundir en acceso abierto estos trabajos inéditos de investigación, por tanto, no se trata de una forma de edición o publicación por los canales tradicionales a cuyas publicaciones se le asigna un ISBN.

Los autores que reclaman a veces un ISBN para sus tesis depositadas en el repositorio, lo hacen ante la exigencia de la mayoría de las convocatorias de las agencias de evaluación académica para que las publicaciones aportadas tengan ISBN.

En este sentido la Agencia del ISBN está llevando a cabo, a través de su página web, una campaña de información para autores, editores y agencias de evaluación sobre lo qué es y no es

el ISBN, quién puede solicitarlo y para qué obras puede hacerlo, además de recriminar a las agencias de evaluación de entorpecer el trabajo de la Agencia del ISBN:

“la exigencia de la mayoría de las convocatorias de las agencias de evaluación académica de que las publicaciones aportadas tengan ISBN pervierte el propósito de la norma y entorpece el funcionamiento de la Agencia del ISBN”, http://agenciaisbn.es/web/view_noticia.php?id=7.

II.1.4. Diseminación de las tesis doctorales de Gredos: la interoperabilidad

Para que los investigadores y las instituciones consigan visibilidad es muy importante que los repositorios institucionales aseguren la interoperabilidad y la preservación digital de sus contenidos con el fin de potenciar la difusión, visibilidad, accesibilidad y uso a largo plazo de los mismos.

Schöpfel (2013) argumenta que para incrementar el descubrimiento y la disponibilidad de las tesis electrónicas de los repositorios es necesario que estos estén conectados entre sí formando redes de trabajo de trabajo del tipo de la NLDTD o de TDR.

Se viene defendiendo a lo largo de esta investigación, el nivel de visibilidad del contenido de un repositorio está directamente relacionado con la interoperabilidad del repositorio, tema que ya se ha tratado en el módulo anterior y, que como se expuso en el mismo, en el campo de los RIs el término interoperabilidad representa un método específico de interoperabilidad como es la recolección de metadatos a través de OAI-PMH (Hakimjavadi & Masrek, 2013).

Todos los repositorios desde el más grande al más pequeño necesitan ser interoperables para que su contenido forme parte del conocimiento global en acceso abierto y de esta forma incrementa la visibilidad de sus contenidos a través de Internet.

En el caso de Gredos la interoperabilidad se aplica a todos los contenidos del repositorio incluidas las tesis doctorales. Como se ha visto anteriormente al hablar del primer informe de

COAR (2012) las áreas en las que más se centra la implementación de herramientas para alcanzar la máxima interoperabilidad de los repositorios son siete: recolección de metadatos, redes de trabajo repositorios, uso de estadísticas, sistemas de intercambio y transferencia de información, sistemas de identificación única de autores, identificadores persistentes, gestión de los distintos componentes de los objetos.

El requisito mínimo para que un repositorio sea conforme a la OAI es que exponga sus metadatos en formato Dublin Core simple y cumpla con el protocolo OAI-PMH. Los portales de tesis usan estos estándares para recolectar la comunidad de tesis doctorales de Gredos.

Otra área importante para investigar sobre la interoperabilidad del repositorio Gredos son las estadísticas que pueden utilizarse para dar a conocer el estado del repositorio y hacer posible su *benchmarking*.

Los identificadores persistentes (URI, DOI) han sido también un tema de interoperabilidad resuelto en el repositorio Gredos desde su puesta en marcha. El sistema *Handle* utilizado por el repositorio está diseñado para designar con una URI a cada uno de los ítems del repositorio. Se trata de un sistema que permite la asignación de identificadores persistentes a los recursos de información u objetos digitales existentes en Internet (artículos, revistas, imágenes, etc.). Desarrollado por *Corporation for National Research Initiatives* (CNRI), se estructura en dos partes:

- Prefijo (Prefix): identifica al productor del identificador (universidad, editorial, revista, etc.). En el caso de Gredos es el "10366".
- Sufijo (Suffix): identifica a cada uno de los documentos u obras digitales (artículo, libro, capítulo, etc.).

La suma del prefijo y el sufijo conforma el identificador persistente, en este caso llamado *handle*. Por ejemplo, el *handle* de la comunidad de tesis doctorales del repositorio Gredos sería: <http://gredos.usal.es/handle/10366/4756>.

La colección de tesis doctorales, al ser difundida a través de Gredos, puede ser consultada desde muchos sitios en la Web, de esta forma incrementa el número de visitas y descargas de su contenido. Actualmente los metadatos de esta colección son recolectados por recolectores temáticos o portales específicos de tesis: TDR, *DART-Europe*, y OATD. Además de estos sitios, las tesis de Gredos se difunden a través de los portales y servicios que recolectan el resto de las colecciones del repositorio: Hispana, Europea, BASE, Recolecta y OpenAire entre otros. Además son indexadas por Google y Google Scholar. En la Figura 29 se muestra gráficamente la difusión de la comunidad de tesis doctorales del repositorio Gredos fruto de la interoperabilidad del repositorio.

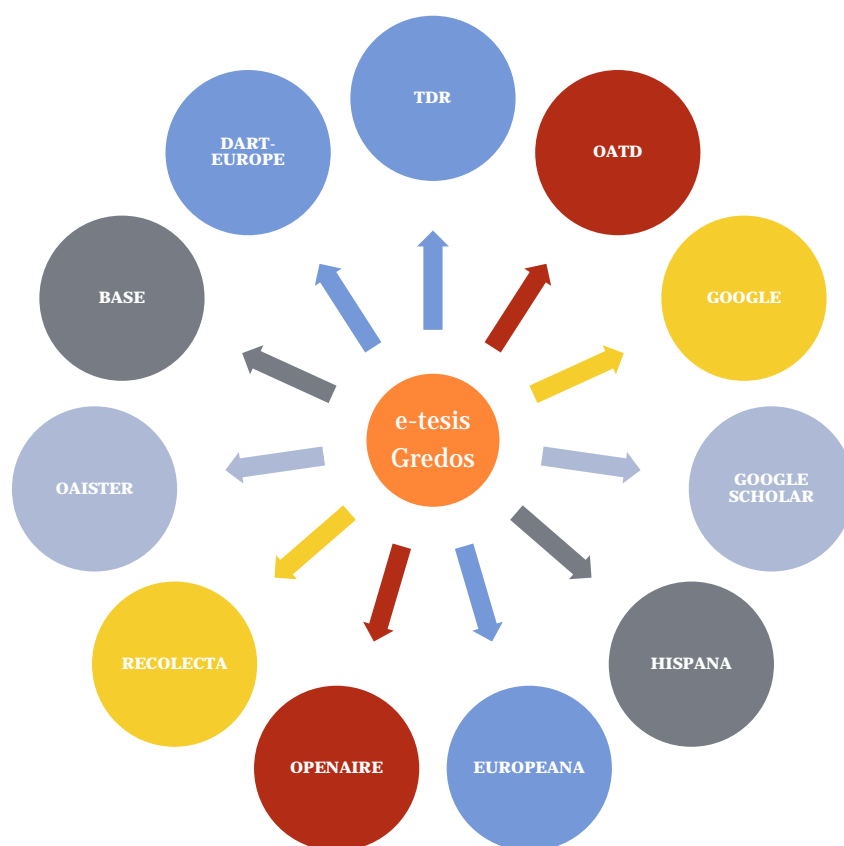


Figura 29. Diseminación del contenido del repositorio Gredos. Fuente: Elaboración propia.

Los resultados de visibilidad del repositorio Gredos son atribuidos al uso de herramientas de interoperabilidad y a la implementación de esquemas de descripción compatibles con las redes de repositorios, con los recolectores, con *Google Scholar* y otros portales de búsqueda (Ferrerías-Fernández, García-Peñalvo, et al., 2015; Ríos-Hilario et al., 2013; Ríos-Hilario, Ferreras-Fernández, & Martín-Campo, 2014; Ríos-Hilario et al., 2012).

También se considera que el desarrollo de buenas prácticas de metadatos de acuerdo a estándares asegura la calidad de los registros de metadatos, incrementando la posibilidad de descubrimiento de las fuentes, incrementando la interoperabilidad de las colecciones de Gredos, permitiendo la recolección por los agregadores de contenido como DRIVER, OpenAire y Europeana, y permitiendo la gestión de la preservación a largo tiempo (García-Peñalvo, Merlo-Vega, et al., 2010).

Para describir las tesis en Gredos se utilizan plantillas de metadatos *Qualified Dublin Core* de *DSpace*, previamente establecidas y de acuerdo a las recomendaciones y directrices de las redes de trabajo de tesis. En relación a este tema, se elaboraron manuales y documentos de trabajo para la descripción con *Dublin Core* de las diferentes tipologías documentales recogidas en Gredos (Ferrerías-Fernández, 2008, 2010b).

Otro logro de interoperabilidad que contribuyó al aumento de la visibilidad de Gredos es la contribución a Europeana (Ferrerías-Fernández, 2013a) con la totalidad de su contenido, gracias a la implementación del esquema de metadatos *Europeana Semantic Elements* (ESE) para todo el contenido de Gredos⁴⁹.

Una evidencia más de visibilidad gracias a la interoperabilidad son los datos proporcionados por Google Analytics, concretamente a través de *referral traffic*, que muestra las fuentes desde las que se accede al repositorio Gredos. Este hecho prueba que la interoperabilidad del repositorio

⁴⁹ Gredos en Europeana (18 de marzo de 2011) [Mensaje de blog] . Recuperado de: <http://diarium.usal.es/gredos/2011/03/18/gredos-en-europeana/>

con otros sitios contribuye a incrementar las visitas: *Europeana*, *Hispana*, *Google Scholar*, *TDR*, *DART-Europe*, *BASE*, *OATD*, y otros (Figura 30).

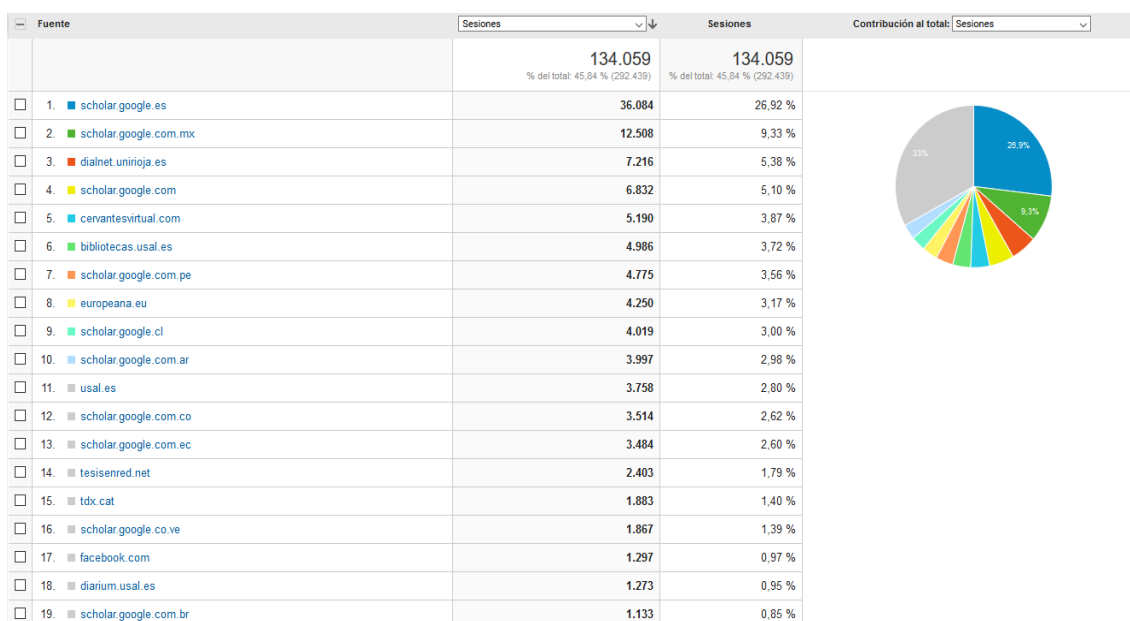


Figura 30. Informe mensual de Gredos de Google Analytics. Fuente: Google Analytics.

A través de estos datos se observa que gracias a la interoperabilidad del repositorio Gredos, el contenido del mismo se disemina a través de recolectores, agregadores y portales, alcanzando así la máxima visibilidad y uso.

Capítulo II.2. Material y métodos

II.2.1. Muestra

La muestra de población seleccionada se compone por una parte de todas las tesis leídas en la Universidad de Salamanca en los dos cursos previos (2006-2007 y 2007-2008) al mandato institucional de acceso abierto de 2008 y por otra parte de todas las tesis leídas en la Universidad de Salamanca en los cursos 2008-2009, 2009-2010 y 2010-2011 que estaban sujetas a mandato, siendo preceptivo el depósito de las mismas en el repositorio de acceso abierto.

A partir del curso 2011-2012 entra en vigor el Real Decreto 99/2011, de 28 de enero (BOE 10/2/2011) que regula los estudios de Doctorado. Aunque la Universidad de Salamanca aplica el artículo 14.5 del Real Decreto donde se regula el depósito de las tesis en el repositorio de acceso abierto, también aplica de forma *ad hoc* la excepción del mismo artículo 14.6, solicitando a los autores de las tesis su autorización para el depósito de las mismas en el repositorio.

De esta forma, el mandato de la Universidad de Salamanca de 2008 se supone derogado por la aplicación del Real Decreto 99/2011. Así, la aplicación del artículo 14.6 supone que no todas las tesis leídas en la Universidad de Salamanca a partir del curso 2011-2012 están disponibles en su versión íntegra en acceso abierto a través del repositorio Gredos.

Por ello, para la muestra se ha escogido el período de acceso abierto total sin condiciones, regulado por el mandato de 2008. Otra razón por la que se ha seleccionado este período de años (2008-2011) ha sido con la finalidad de que en la toma de datos, realizada unos años después en agosto 2014, sobre uso (visitas y descargas) e impacto (citas) hubiera suficiente información para realizar el análisis. De acuerdo a Brody, Harnad y Carr el impacto de citación de un artículo solo puede medirse varios años después de ser publicado (Tim Brody et al., 2006).

En esta investigación se ha considerado que para realizar el análisis de uso e impacto de las tesis en el repositorio Gredos, estas deberían llevar publicadas en acceso abierto varios años, para el caso tratado entre 2 y 3 años, con el fin de obtener datos consistentes para el análisis.

II.2.2. Selección del material

Las fuentes de datos utilizadas en la selección del material han sido:

1. El repositorio Gredos.
2. La base de datos TESEO.

La relación de tesis doctorales leídas en la Universidad de Salamanca constituye la base de datos en la que se sustenta esta investigación.

Se utilizó la base de datos TESEO para localizar todas las tesis leídas en la Universidad de Salamanca en el período estudiado (2006-2011). Una primera búsqueda en TESEO consistió en localizar todas las tesis recuperándolas por el criterio “cursos académicos”. Esta primera búsqueda arrojó un total de 1.182 tesis leídas en la Universidad de Salamanca durante 5 cursos académicos.

La base de datos TESEO⁵⁰, editada por el Ministerio de Educación, permite recuperar información sobre las tesis doctorales defendidas en las universidades españolas desde 1976. La información proporcionada por este recurso en línea ha fomentado el desarrollo de numerosos estudios bibliométricos y cienciométricos sobre el estado de la cuestión en la investigación española sobre diversas áreas de conocimiento y especialidades.

⁵⁰ TESEO, <https://www.educacion.gob.es/teseo>

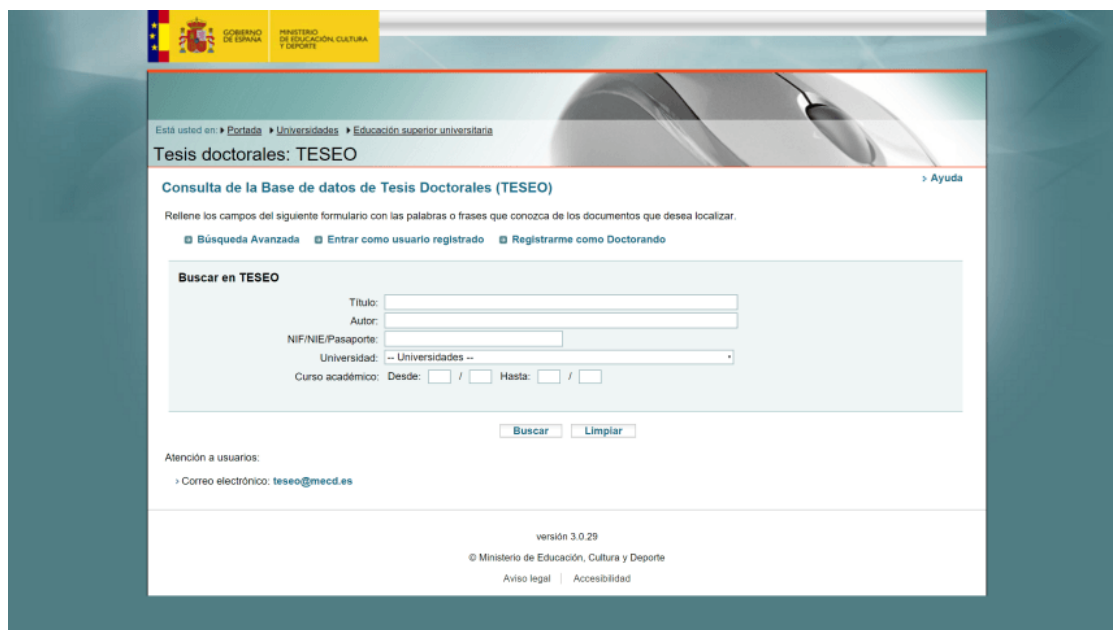


Figura 31. Base de datos TESEO. Pantalla de consulta. Fuente: TESEO.

Se realizaron búsquedas por “Universidad de Salamanca” limitando por curso académico. Se depuró la búsqueda porque los resultados presentaban muchos errores, tales como fichas duplicadas y duplicidad de autores. El manejo de los resultados de búsqueda en esta base de datos es muy difícil: la lista de títulos resultante de la búsqueda no se puede ordenar con ningún criterio y no permite la exportación de los resultados para poder manejarlos y analizarlos.

Los datos de cada una de las 1.182 fichas de tesis se codificaron manualmente en ficheros *Excel*. Los datos extraídos de cada tesis fueron: autor, título, directores, departamento, fecha de lectura. De esta forma las tesis se clasificaron por curso, y por área de conocimiento a partir del dato “departamento”.

Las cifras obtenidas de TESEO sobre el número de tesis leídas por curso se expresan en la Tabla 32.

Tabla 32

Datos TESEO. N° de tesis de la U. Salamanca por cursos

CURSO	TESEO
2006-2007	220
2007-2008	243
2008-2009	242
2009-2010	222
2010-2011	255
	1182

Nota: Elaboración propia

Según estos datos las tesis que no están en acceso abierto por no estar sometidas a mandato son 463 tesis correspondientes a los cursos 2006-2007 / 2007-2008, mientras que 719 tesis deberían estar en acceso abierto.

Todas las tesis se clasificaron en cinco categorías de acuerdo a las áreas de conocimiento a las que pertenecen los departamentos donde se realizaron las tesis: Ciencias biosanitarias, Ciencias experimentales, Humanidades, Ciencias sociales y Ciencias técnicas (Tabla 33 y Figura 32).

Tabla 33

Clasificación de las tesis doctorales por categorías

CATEGORIAS	OA Mandato	No OA	TOTAL TESIS
Ciencias biosanitarias	214	127	341
Ciencias experimentales	102	63	165
Ciencias sociales	254	180	432
Enseñanzas técnicas	24	16	40
Humanidades	127	77	204
	721	463	1182

Nota: Elaboración propia

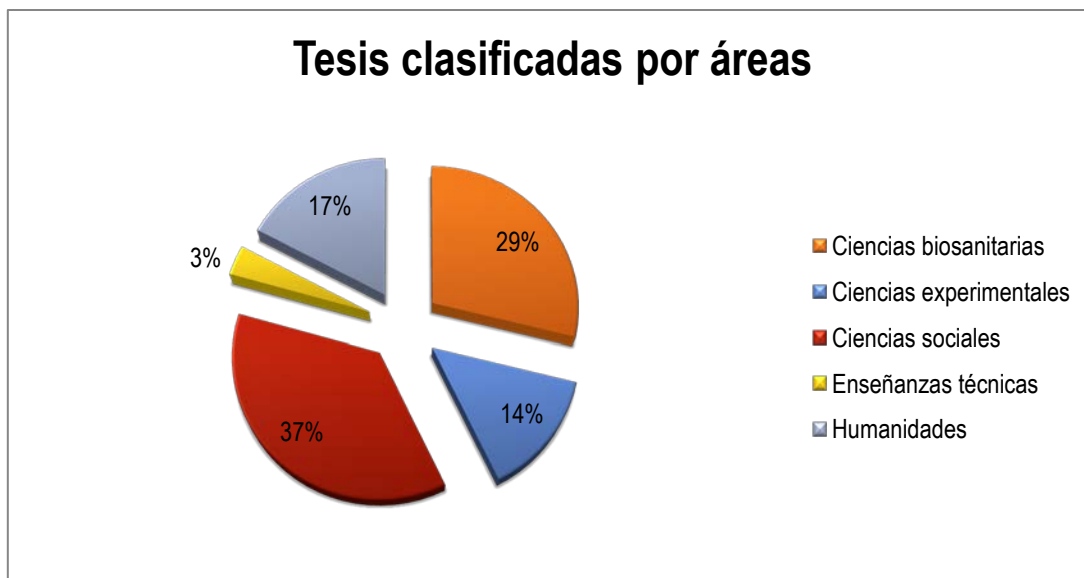


Figura 32. Porcentaje de tesis por áreas de conocimiento. Fuente: Elaboración propia.

A raíz de la publicación del análisis de la base de datos TESEO realizado por Manuel Blázquez Ochando a finales de 2015, <http://mblazquez.es/catalogo-de-tesis-doctorales-espanolas-teseo-disponible-para-su-descarga/>, se realiza una segunda búsqueda en TESEO cambiando el criterio de la búsqueda, en lugar de buscar por “cursos académicos” se buscó por años naturales, ya que según el análisis de Blázquez Ochando se detectaron muchas fichas de tesis doctorales, en concreto 7.857, sin el dato de la fecha de lectura de la tesis, por lo que la clasificación por años que utilizó este autor ha parecido más fiable a la hora de la elaboración de la base de datos de esta investigación.

Blázquez Ochando⁵¹ realizó una descarga completa de la base de datos TESEO que le permitió realizar los primeros análisis y comprobaciones, de los que se han obtenido algunos datos e informaciones interesantes que se describen a continuación.

La base de datos TESEO dispone de 196.442 Tesis Doctorales registradas para las que existen 1.190.508 marcadores, ello implica que 994.066 son duplicados. Se detectan 7.857 fichas de Tesis Doctorales sin el dato de la fecha de lectura. Según el número de tesis doctorales

⁵¹ Catálogo de Tesis Doctorales Españolas Teseo disponible para su descarga (<http://mblazquez.es/catalogo-de-tesis-doctorales-espanolas-teseo-disponible-para-su-descarga/>) [Artículo actualizado y revisado 2015-11-15] Prof. Dr. Manuel Blázquez Ochando

registradas en TESEO, tomando como referencia la fecha de lectura, se observa un balance general de crecimiento positivo. Se puede confirmar que los momentos de mayor crecimiento se concentran en los periodos “1986-1987” con un incremento de +2.260, “1989-1990” con +1.474, “1993-1997” con +2.694 y finalmente 2007-2014 con un diferencial de crecimiento de +4.061 nuevos trabajos registrados.

En cuanto a los datos refereridos exclusivamente a la Universidad de Salamanca, se observa que en el período comprendido entre 1977 a 2015 se leyeron un total de 5.863 tesis doctorales en la Universidad de Salamanca. La evolución por años aparece reflejada en los siguientes gráficos (Figura 33 y figura 34).

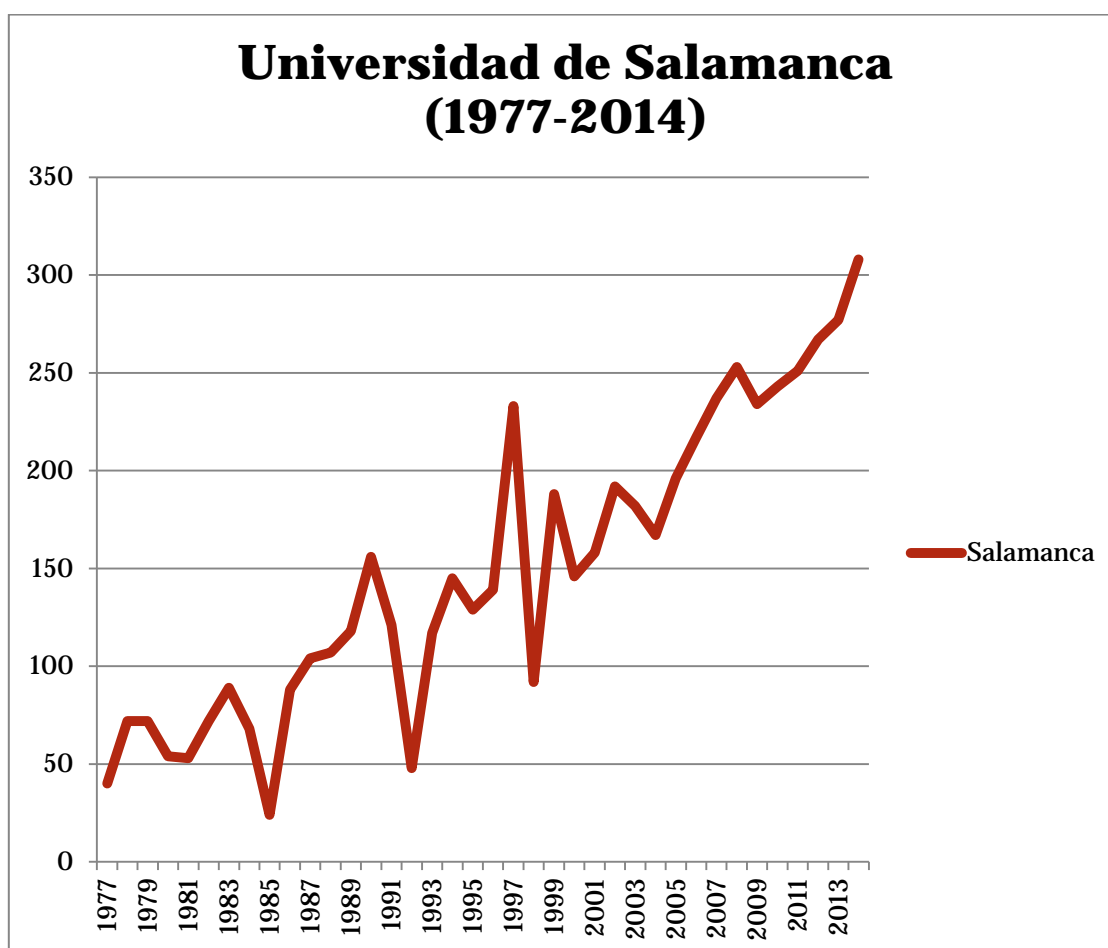


Figura 33. Tesis Doctorales en la Universidad de Salamanca I (1977-2014). Fuente: Modificado de Blázquez Ochando (2015).

Tesis Doctorales en la Universidad de Salamanca (1977-2014)

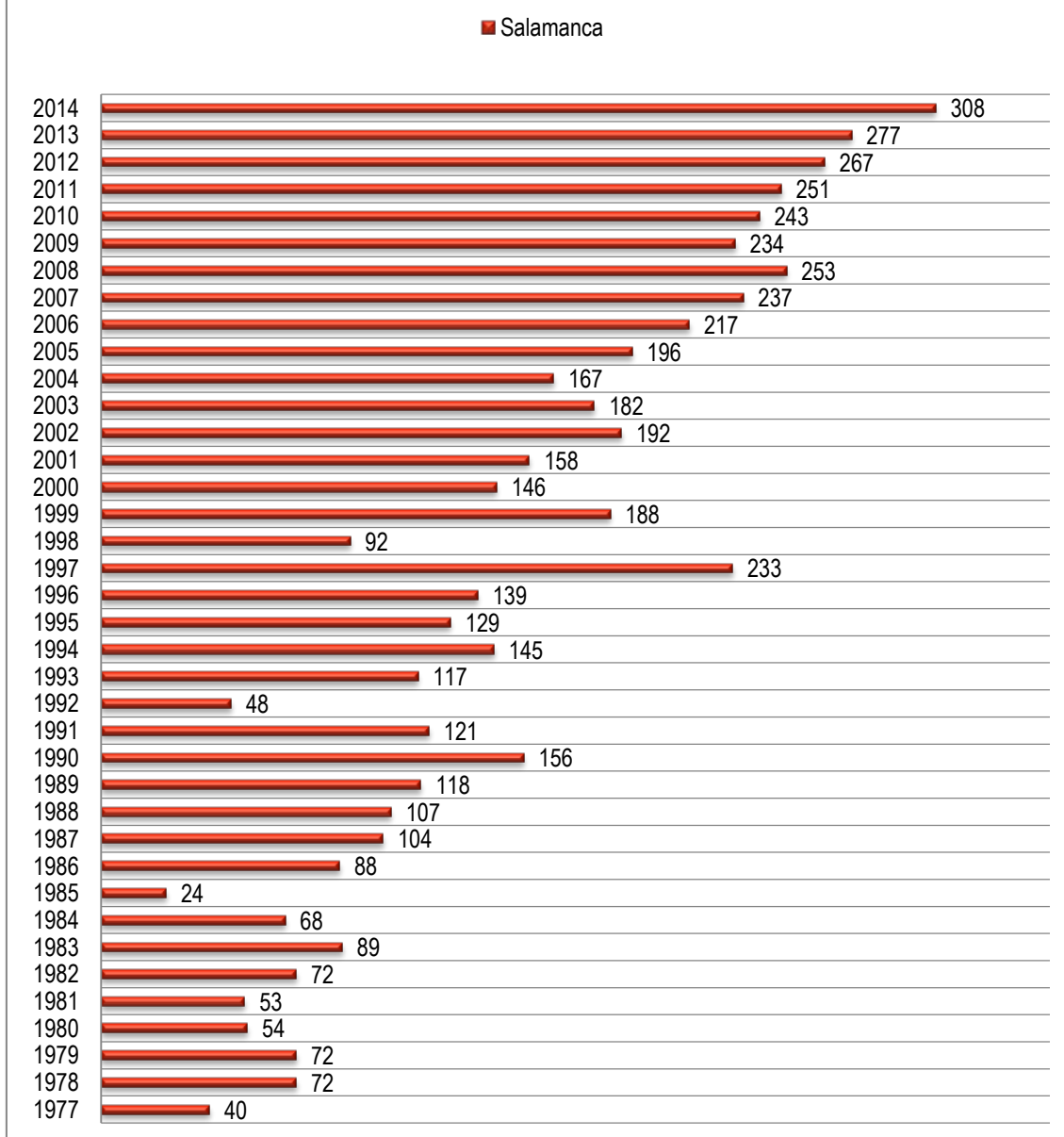


Figura 34. Tesis Doctorales en la Universidad de Salamanca II (1977-2014). Fuente: Modificado de Blázquez Ochando (2015).

Del período referido en esta investigación, años 2006-2011, se han hallado un total de 1.435 tesis doctorales (Tabla 34). Por tanto, existe una diferencia positiva de 253 tesis más, potencialmente en acceso abierto, al extraer los datos por años naturales en lugar de extraerlos por cursos académicos como se hizo en la primera búsqueda.

Tabla 34

Tesis de la Universidad de Salamanca

TESIS U. SALAMANCA, 2006-2011 (n=1435)					
Tesis NO OA (n=454)		Tesis OA (n=981)			
2006	2007	2008	2009	2010	2011
217	237	253	234	243	251

Nota: Elaboración propia

Para realizar el análisis se han utilizado los datos de la segunda búsqueda, de esta forma se han seleccionado la totalidad de las 1.435 tesis doctorales correspondientes al período 2006-2011. El objetivo de la presente tesis doctoral es analizar la visibilidad, uso e impacto de las tesis en acceso abierto depositadas en el repositorio institucional Gredos, por lo que de estas se han seleccionado 863 tesis correspondientes al período 2008-2011 que son las que se encuentran depositadas en el repositorio institucional Gredos.

Para realizar dicho estudio, se creó una base de datos que contuviera los trabajos publicados y citados en dicho periodo.

La extracción de datos desde el repositorio Gredos se realizó mediante la exportación de metadatos de las tesis alojadas en la comunidad *Tesis doctorales*, <http://gredos.usal.es/jspui/handle/10366/4756>, correspondientes a los años 2008-2011. El fichero generado para la exportación de metadatos fue un fichero csv que se gestionó en *Excel* para filtrar los metadatos descriptivos que se utilizarían para el estudio: por una parte se trataba de los mismos datos que habíamos extraído de TESEO tales como autor, título, director, departamento, y fecha de lectura, y por otra parte se añadieron datos como la URI que identificaba en el repositorio cada ítem y los datos estadísticos de cada ítem proporcionados por el repositorio.

Según TESEO son 981 tesis, el número de tesis leídas entre 2008-2011 y que por tanto estarían sujetas a mandato institucional para su depósito en acceso abierto en el repositorio. Sin embargo, el número de tesis depositadas en Gredos de ese período de tiempo es de 863, las cuales serán el objeto de estudio de esta investigación doctoral.

II.2.3. Definición de las variables

Las variables estudiadas son de tipo cualitativo (*áreas o disciplinas analizadas*) y cuantitativo: variables relacionadas con el uso y/o visibilidad (*visitas y descargas*) y con el impacto (*citas*).

La variable "*área o disciplina temática*" consta de 5 categorías (1= CC Biosanitarias; 2= E. Técnicas; 3= CC. Sociales; 4= Humanidades y 5= CC. Experimentales).

En relación a las variables cuantitativas:

- **Visitas:** hace referencia a las visitas recibidas y contabilizadas por el módulo de estadísticas de la Universidad de Tasmania instalado en el repositorio Gredos. Cada ítem o registro del repositorio Gredos muestra las estadísticas de visitas que ha recibido.
- **Descargas:** hace referencia a los datos de descargas contabilizadas, al igual que las visitas, por el módulo de la Universidad de Tasmania instalado en el repositorio Gredos. Cada ítem o registro del repositorio Gredos muestra las estadísticas de descargas que ha tenido. Los datos se extrajeron de forma manual entrando en cada uno de los registros y recogiendo los datos de descargas.
- **Citas o citación:** se han contabilizado a través de *Google Scholar*. Cada registro está monitorizado por la API de *Google Scholar* que permite obtener información sobre las citas recibidas de cada uno de trabajos del repositorio Gredos.

La información de las tres variables se recopiló manualmente en Excel como se muestra en el ejemplo de la Figura 36, por una parte entrando desde Gredos en cada uno de los registros de la

colección tesis doctorales y tomando la información de visitas y descargas recibidas a través del módulo de estadísticas, y por otra parte desde la API de *Google Scholar* instalada en cada uno de los ítems de Gredos se tomó la información sobre las citas recibidas.

ID	Autor	Título	Fecha	Dirección	Area	URI	Visitas	Descarga GS	Citación
1	Abadía Cubillo, Karla	Ejercicio físico como parte de la atención de los par	2010	Cruz Hernández, Juan Jesús Sánchez Domínguez,	Biosanitarias	http://hdl.handle.net/10366/83192	593	6541	yes 0
2	Abollo Jiménez, Fernando	Identificación y caracterización funcional de nuevo	2011	Sánchez García, Isidro Cobaleda Hernández, César	Biosanitarias	http://hdl.handle.net/10366/110637	301	946	yes 0
3	Abreu Ferreira Machado, Carlos Albe	La vertebración del territorio ibérico de la raya duri	2011	Cabero Diéguez, Valentín	Humanidades	http://hdl.handle.net/10366/115563	347	1447	yes 0
4	Agudelo Ramírez, Martín	El problema de la fundamentación filosófica de los	2010	Álvarez Gómez, Mariano	Humanidades	http://hdl.handle.net/10366/76364	283	2163	yes 0
5	Aguilera Izaguirre, Gustavo	Modelos de evaluación de la calidad de las organiz	2010	Nevado-Batalla Moreno, Pedro Tomás	Sociales	http://hdl.handle.net/10366/76368	441	1633	yes 0
6	Airado Cerrato, Carmen	Estudio del sistema zincérgico en el bulbo olfatorio	2010	Alonso Peña, José-Ramón Gómez Rodríguez, Carmela	Biosanitarias	http://hdl.handle.net/10366/76367	167	1394	yes 0
7	Albarrán Diego, Juan	Del fotoconceptualismo al fototableau. Fotografía, j	2011	Panera Cuevas, Francisco Javier	Humanidades	http://hdl.handle.net/10366/110638	719	3038	yes 5
8	Albuquerque, Bruna María Jacques F	Subcontratación y precarización del trabajo. Un est	2011	Baz Rodríguez, Jesús Cabero Morán, Enrique	Sociales	http://hdl.handle.net/10366/110697	303	6297	yes 0
9	Alburquerque Sendín, Francisco	Estudio comparativo intermetodológico de la comp	2008	Santos del Rey, Miguel Martín Vallejo, Francisco Javier	Biosanitarias	http://hdl.handle.net/10366/17593	516	4482	yes 2
10	Alcoeba Sánchez, Miguel	Estudio de polimorfismos genéticos en la evolución	2010	González Díaz, Marcos García Sanz, Ramón	Biosanitarias	http://hdl.handle.net/10366/76378	180	4491	yes 0
11	Aldana Bermúdez, Eliécer	Comprensión del concepto de integración definida i	2011	González Astudillo, María Teresa	Sociales	http://hdl.handle.net/10366/83204	318	7500	yes 8
12	Aliste Santos, Tomás Javier	La motivación de las resoluciones judiciales	2010	Bujosa Vadell, Lorenzo Mateo	Sociales	http://hdl.handle.net/10366/76379	Retirado	Retirado	no

Figura 35. Ejemplo de la base de datos construida en Excel. Fuente: Elaboración propia.

Toda la información de las tres variables se recopiló durante el mes de agosto de 2014.

Para facilitar la realización de algunos de los análisis efectuados, “*el año*”, que hace referencia al año de lectura y depósito en el repositorio institucional, se recodificó con valores que van del 1 al 4 (1= año 2008, 2=año 2009, 3=año 2010 y 4=año 2011). La variable “*citas*” también se consideró categorizada: 0 = ninguna cita y 1 = al menos una cita.

II.2.4. Análisis y métodos estadísticos

El acceso abierto a la información ha dado lugar a una multiplicación de los datos que se pueden utilizar para llevar a cabo análisis estadísticos sobre el impacto de la publicaciones científicas, preparando así el camino para nuevos modelos de métricas para la comunicación académica (Bernal, 2013; Bernal & Pemau-Alonso, 2010).

Las fuentes utilizadas para extraer los datos con los que realizar los distintos análisis han sido:

1. Bases de datos multidisciplinares: *Web of Science* (WOS) y Scopus.
2. La colección de Tesis Doctorales del repositorio Gredos.

3. El buscador científico *Google Scholar*.

II.2.4.1. Análisis de las citas en las bases de datos WOS y Scopus

En primer lugar, se realizó un análisis basado en las bases de datos *Web of Science* (WOS) y Scopus con el fin de localizar las tesis de la Universidad de Salamanca del período estudiado que hubiesen recibido citas. Se trataba de averiguar los datos de citación tanto de las tesis que no estaban depositadas en acceso abierto como de las tesis que sí lo estuvieran. Los datos resultantes podrían ofrecer una comparación entre el impacto de tesis no OA y de las tesis OA.

Posteriormente, se abandonó el análisis en Scopus ya que no se pudo replicar la búsqueda empleada en WOS para la extracción de datos.

La base de datos WOS indexa muchos tipos de documentos como por ejemplo las fuentes de los ítems, entre las que generalmente solo se utilizan los artículos, las notas de investigación y las revisiones de artículos para hacer estudios bibliométricos.

Para identificar las referencias a las tesis doctorales entre todo el material citado, se llevó a cabo un proceso de búsqueda a través de la opción “Búsqueda de referencia citada” utilizando la fórmula “thesis* U Salamanca” en el campo “Trabajo citado”, combinando con “Año de cita” (2006-2011) y limitando por el período de tiempo entre 2006 hasta 2014, año en el que se recogieron los datos de uso e impacto de toda la muestra de esta investigación. Por su parte Scopus no permite buscar por trabajo citado en un campo específico para ello, por lo que la búsqueda de las tesis citadas habría que realizarla por los datos de autor y título de cada una de las 1.435 tesis doctorales correspondientes al período de tiempo estudiado. Este hecho motivó que en este punto se abandonase esta opción para extraer datos y solo se utilizase WOS.

La Figura 36 muestra las búsquedas realizadas en WOS.

Trabajo citado: (thesis* U Salamanca) AND Año de cita: (2006 OR 2007 OR 2008 OR 2009 OR 2010 OR 2011)

Período de tiempo=2006-2014

The screenshot shows the Web of Science search interface. At the top, there is a navigation bar with the 'WEB OF SCIENCE' logo and the Thomson Reuters logo. Below this, there is a search bar with the text 'tesis* U Salamanca' and a dropdown menu for 'Trabajo citado'. There are also dropdown menus for 'Año(s) de cita' and a 'Buscar' button. The interface includes instructions for searching and a note about maintenance. The 'PERÍODO DE TIEMPO' section is set to 'Desde 2006 hasta 2014'.

Figura 36. Búsqueda de tesis de la Universidad de Salamanca de 2006-2007. Fuente: WOS.

WOS permiten exportar los resultados de las búsquedas en tablas de datos que se pueden gestionar en Excel para hacer control de calidad ampliando datos y corrigiendo errores.

Los resultados obtenidos en WOS se comentan más adelante en el módulo 3 correspondiente a Resultados.

II.2.4.2. Análisis de visibilidad, uso y citación de las tesis de Gredos

Como se ha indicado en el epígrafe 2.3. sobre la definición de las variables, los datos sobre las mismas (nº de visitas, nº de descargas y nº de citas) se extrajeron de Gredos y de *Google Scholar*, y se recopilaron en *Excel*.

Para la obtención y análisis de datos se utilizó el programa *IBM SPSS para Windows*, Version 23.0 (IBM Corp, Armonk, NY), y la hoja de cálculo *Excel 2007* que facilitan los cálculos matemáticos y la construcción de tablas y figuras.

El control de calidad de la información se realizó a través de dobles tablas, corrigiendo mediante la consulta con los originales los errores que se detectaron.

Se realizó un análisis descriptivo y comparativo de las variables objeto de estudio.

Las variables cualitativas se han descrito con sus frecuencias absolutas y porcentajes, representándose mediante la utilización de tablas y gráficos. Para determinar su asociación se empleó el test de χ^2 de Pearson.

En estadística, la prueba χ^2 de Pearson se considera una prueba no paramétrica que mide la discrepancia entre una distribución observada y otra teórica (bondad de ajuste), indicando en qué medida las diferencias existentes entre ambas, de haberlas, se deben al azar en el contraste de hipótesis. También se utiliza para probar la independencia de dos variables entre sí, mediante la presentación de los datos en tablas de contingencia.

Las variables continuas han sido expresadas como media \pm desviación estándar y como mediana [rango intercuartílico]. Las pruebas de distribución de normalidad se hicieron con el test de Kolmogórov-Smirnov.

En estadística, la prueba de Kolmogórov-Smirnov (también prueba K-S) es una prueba no paramétrica que determina la bondad de ajuste de dos distribuciones de probabilidad entre sí. La prueba Kolmogórov-Smirnov es más sensible a los valores cercanos a la mediana que a los extremos de la distribución.

Mediante el coeficiente de correlación de Spearman se estimó la relación entre las variables cuantitativas del estudio. En estadística, el coeficiente de correlación de Spearman, ρ (rho) es una medida de la correlación (la asociación o interdependencia) entre dos variables aleatorias continuas. Para calcular ρ , los datos son ordenados y reemplazados por su respectivo orden.

Para analizar la comparación de los indicadores de visibilidad e impacto según las áreas se utilizó el test de Kruskal-Wallis, ya que es el método más adecuado para comparar poblaciones cuyas distribuciones no son normales.

En estadística, la prueba de Kruskal-Wallis es un método no paramétrico para probar si un grupo de datos proviene de la misma población. Ya que es una prueba no paramétrica, la prueba de Kruskal-Wallis no asume normalidad en los datos aunque asume, bajo la hipótesis nula, que los datos vienen de la misma distribución.

Para Kruskal-Wallis, las hipótesis son:

- H_0 : todas las medianas de las poblaciones son iguales.
- H_1 : todas las medianas de las poblaciones no son iguales.

$$H = \frac{12}{N(N-1)} \sum_{i=1}^k \left(\frac{R_i^2}{n_i} \right) - 3(N+1)$$

$$H = \frac{\frac{12}{N(N-1)} \sum_{i=1}^k \left(\frac{R_i^2}{n_i} \right) - 3(N+1)}{1 - \frac{\sum_{i=1}^g (t_i^2 - t_i)}{N^3 - N}}$$

Donde g denota el número de grupos de ligaduras y t_i el total de números de orden ligados en el i -ésimo grupo.

Al ser el tamaño muestral mayor a 5, el estadístico H sigue una distribución χ^2 con $k-1$ grados de libertad.

En la prueba de Kruskal-Wallis la hipótesis alternativa es que no todos los grupos tienen la misma distribución, esto es, que en al menos en dos grupos hay diferencias.

Para averiguar entre qué par de grupos se han encontrado diferencias se realizaron comparaciones por pares (*post-hoc*) con el test Dunn.

Módulo III. Resultados y Discusión

En el módulo 3 se presentan los resultados obtenidos y el análisis y discusión de los mismos.

Este módulo consta de 2 capítulos. En el capítulo III.1, se muestran los resultados de los distintos análisis llevados a cabo. En el capítulo III.2, se desarrolla la discusión sobre dichos resultados que conducirán a las conclusiones del siguiente módulo.

La estructura y contenidos del módulo es la siguiente:

Módulo III. Resultados y discusión

Capítulo III.1. Resultados

III.1.1. Resultados del análisis de citación en WOS

III.1.2. Resultados del análisis de visibilidad, uso y citación de las tesis de Gredos

Capítulo III.2. Discusión

III.2.1. Cuestiones previas

III. 2.2. WOS como herramienta de búsqueda de tesis citadas

III. 2.3. Medición de la visibilidad, el uso y el impacto de las tesis doctorales en acceso abierto

III.2.4. Google Scholar como herramienta alométrica para la identificación de citas

Capítulo III.1. Resultados

El total de tesis doctorales defendidas en la Universidad de Salamanca durante el período 2006-2011 fue de 1.435, de las cuales 863 (60,14%; IC95% 57,61-62,67) están en acceso abierto y 572 (39,86%; IC95% 37,33-42,39) no lo están.

III.1.1. Resultados del análisis de citación en *Web of Science*

Según se indica en la Tabla 35 del total de tesis doctorales de la Universidad de Salamanca defendidas en el período de tiempo analizado (1.435 tesis), 66 tesis obtuvieron alguna cita en 77 trabajos (Figura 37) según la base de datos WOS.

Tabla 35

Tesis doctorales citadas por años y nº total de citas por años

Año	Tesis USAL	Tesis citadas	Nº total citas
2006	217	13	20
2007	237	12	16
2008	253	18	18
2009	234	9	9
2010	243	5	7
2011	251	9	10
	1.435	66	80

Nota: Elaboración propia

ÍNDICE DE REFERENCIAS CITADAS									
Referencias: 1 - 50 de 66									
<input type="button" value="Seleccionar página"/> <input type="button" value="Seleccionar todo*"/> <input type="button" value="Borrar todo"/> <input type="button" value="Finalizar búsqueda"/>									
Seleccionar	Autor citado	Trabajo citado [MOSTRAR TÍTULOS COMPLETOS]	Año	Volumen	Número	Página	Identificador	Artículos en que se cita **	Ver registro
<input checked="" type="checkbox"/>	ALVAREZ C	THESIS U SALAMANCA	2006					2	
<input checked="" type="checkbox"/>	Alvarez Fernandez, E.	THESIS U SALAMANCA	2006					1	
<input checked="" type="checkbox"/>	ARENZANA FJ	THESIS U SALAMANCA	2006					1	
<input checked="" type="checkbox"/>	ARROYO MA	THESIS U SALAMANCA	2009			15266		1	
<input checked="" type="checkbox"/>	AYLLON J	THESIS U SALAMANCA	2009					1	
<input checked="" type="checkbox"/>	BRENES BARAHONA, AMELIA	THESIS U SALAMANCA	2010					1	
<input checked="" type="checkbox"/>	Brenes, Diego	THESIS U SALAMANCA	2011					1	
<input checked="" type="checkbox"/>	BUSTOS A	THESIS U SALAMANCA	2009					1	
<input checked="" type="checkbox"/>	Cachaza Gianzo, E. M.	THESIS U SALAMANCA	2011					1	
<input checked="" type="checkbox"/>	CADAVID, M.	THESIS U SALAMANCA	2008					1	
<input checked="" type="checkbox"/>	Carrasco, A.	THESIS U SALAMANCA	2011					2	
<input checked="" type="checkbox"/>	Cid, E.	THESIS U SALAMANCA	2006					1	
<input checked="" type="checkbox"/>	CILLEROS MM	THESIS U SALAMANCA	2009					1	
<input checked="" type="checkbox"/>	CIVIDANES JL	THESIS U SALAMANCA	2006					1	

Figura 37. Índice de referencias citadas de WOS. Fuente: Elaboración propia.

Se ha podido comprobar que los datos ofrecidos por WOS de las tesis doctorales citadas contienen muchos errores que en algunos casos hacen imposible la identificación de los trabajos. Los errores son del tipo: nombre de autores incompleto, no aparecen los títulos de las tesis, algunos de los trabajos identificados como tesis no son en realidad tesis doctorales sino otro tipo de trabajo. En 27 casos ha sido imposible la identificación del trabajo porque carecían de título.

Por tanto, los resultados obtenidos son irrelevantes a la hora de hacer un análisis de citación de la totalidad de las tesis doctorales de la Universidad de Salamanca. Lo que impide realizar un análisis comparativo entre las tesis NO-OA y las tesis OA.

De acuerdo a los datos de citación ofrecidos por WOS, del período comprendido entre 2006 y 2007 se citan 12 tesis doctorales que están publicadas (6 pertenecientes a 2006 y 6 pertenecientes a 2007); mientras que del periodo comprendido entre 2008 y 2011 se citan 19 tesis doctorales que se encuentran en acceso abierto a través de Gredos.

III.1.2. Resultados del análisis de visibilidad, uso y citación de las tesis de

Gredos

Para este estudio se ha partido de una muestra compuesta por 863 tesis leídas en la Universidad de Salamanca en el periodo comprendido entre los cursos académicos 2008-2011 (años 2008, 2009, 2010 y 2011) y que se encuentran depositadas en el repositorio institucional Gredos.

De estas, 48 (5,56%; IC95%) fueron retiradas, en distintos periodos de tiempo, a petición de los autores quedando, en el momento actual, un total de 815 tesis depositadas en el repositorio institucional, lo que supone un 94,44% del total para el período estudiado (Figura 38), siendo todas ellas indexadas en el buscador *Google Scholar*.

Tesis Gredos (2008-2011)

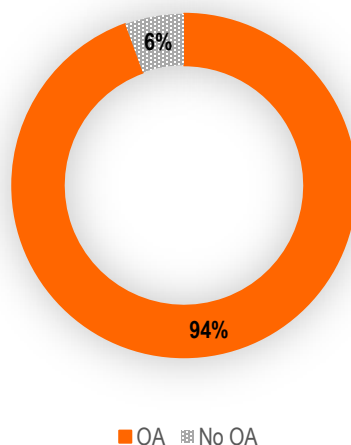


Figura 38. Tesis OA USAL. Fuente: Elaboración propia

Distribución de las tesis doctorales

En relación a la distribución de las tesis doctorales por áreas (véase Tabla 36 y Figura 39) se observa que las áreas con mayor volumen de tesis doctorales presentadas en estos años son Ciencias Sociales (309) y Ciencias Biosanitarias (249).

Tabla 36

Distribución de las tesis doctorales por áreas de conocimiento

Áreas	Frecuencias	Porcentajes	IC95%
CC. Biosanitarias	249	28,90	25,83-31,88
E. Técnicas	32	3,70	2,45-4,97
CC. Sociales	309	35,80	32,61-39,00
Humanidades	156	18,10	15,51-20,64
CC. Experimentales	117	13,60	11,27-15,84
Total	863	100,00	--

Nota: Elaboración propia

En cuanto a la distribución de las tesis doctorales por años (véase Tabla 37 y Figura 39), los años con mayor volumen de depósito en acceso abierto son los años 2010 y 2011.

Tabla 37

Distribución de las tesis doctorales por años

Años	Frecuencias	Porcentajes	IC95%
2008	152	17,60	15,07-20,15
2009	155	18,00	15,40-20,52
2010	297	34,40	31,25-37,58
2011	259	30,00	26,95-33,07
Total	863	100,00	--

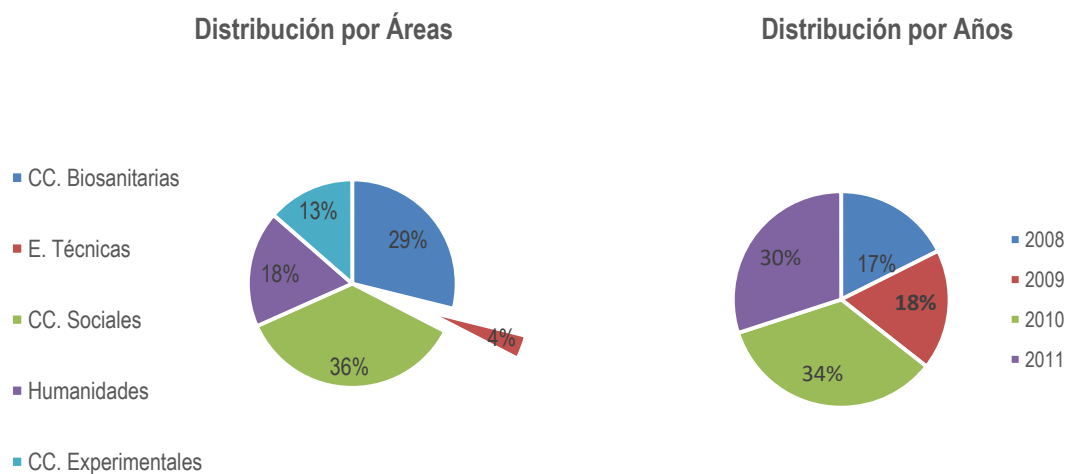


Figura 39. Distribución de las visitas, descargas y citas por áreas y años. Fuente: Elaboración propia

Medias de las variables (visitas, descargas y citas)

A continuación, se presentan las medias de las visitas (Figura 40), descargas (Figura 41) y citas realizadas (Figura 42) con sus respectivos errores estándar. Las barras de error se utilizan con el

fin de indicar el error que se estima en una medida; es decir, una barra de error indica la incertidumbre de un valor.

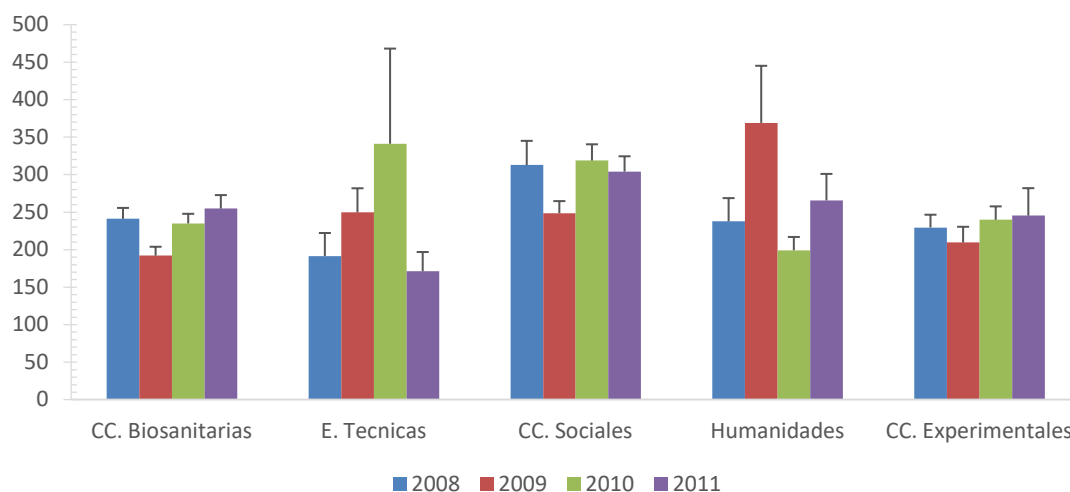


Figura 40. Representación de los valores medios y error estándar de las visitas por áreas y años. Fuente: Elaboración propia

En el gráfico de la Figura 40 se observa la media de visitas recibidas por las tesis en cada una de las áreas y agrupadas por años. En general, reciben más visitas las tesis del área de Ciencias Sociales y destaca el aumento de la media de las visitas en el año 2010 el área de Enseñanzas técnicas, y en el año 2009 el área de Humanidades.

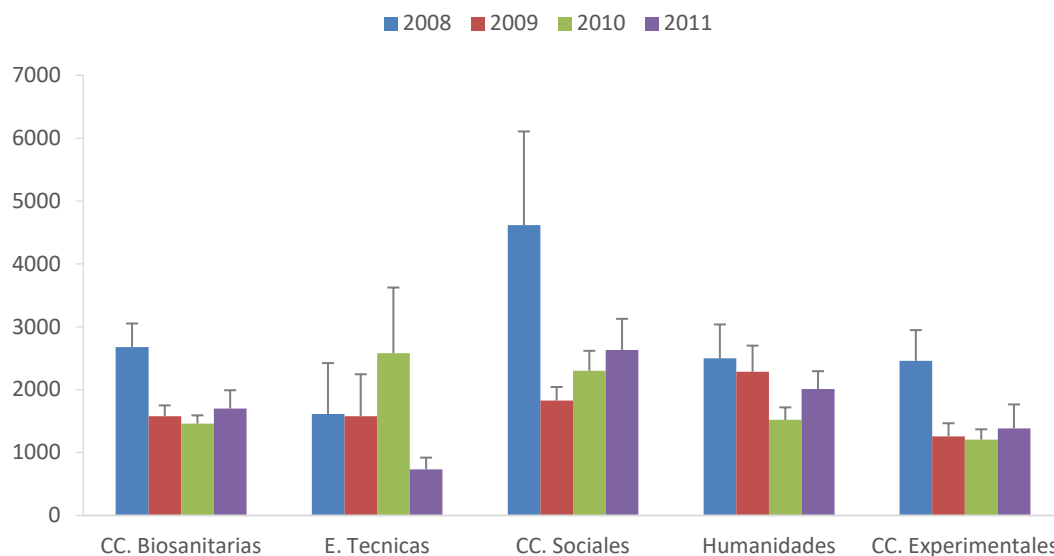


Figura 41. Representación de los valores medios y error estándar de las descargas por áreas y años. Fuente: Elaboración propia

En el gráfico de la Figura 41 se observa la media de descargas recibidas por las tesis en cada una de las áreas y agrupadas por años. En general, la media de las descargas es muy similar en todas las áreas de conocimiento en los años estudiados. Se produce una distorsión en el área de Ciencias Sociales en el año 2008 en el que se eleva el número de descargas recibidas.

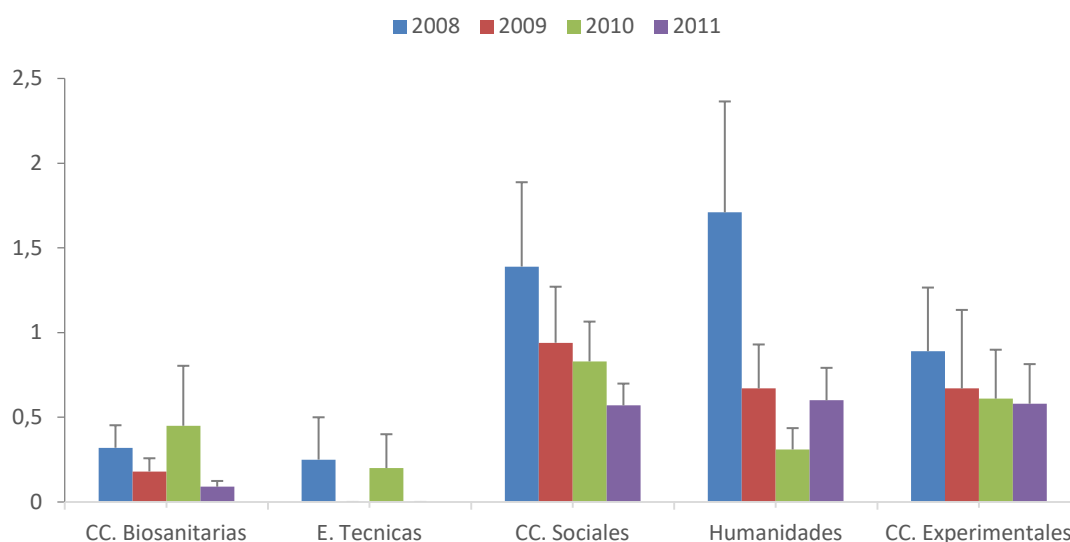


Figura 42. Representación de los valores medios y error estándar de las citas por áreas y años. Fuente: Elaboración propia

En cuanto a la media de citas recibidas, en la Figura 42 se muestra claramente, a través de la representación de los valores medios, que son las áreas de Ciencias Sociales, Humanidades y Experimentales las que reciben más citas. Destaca en el año 2008 el aumento de citas en Ciencias Sociales y Humanidades.

Indicadores de visibilidad (visitas y descargas) e impacto

De los resultados obtenidos, se han encontrado diferencias estadísticamente significativas entre las distintas variables analizadas (indicadores de visibilidad e impacto) en las distintas áreas. En la Tabla 38 se observan valores medios para cada uno de los indicadores.

Tras realizar el análisis de contrastes *post-hoc* se observó que existían diferencias estadísticamente significativas para la variable "visitas" entre las áreas de Enseñanzas Técnicas y Ciencias Sociales, Ciencias Biosanitarias y Ciencias Sociales, Ciencias Experimentales y

Ciencias Sociales y finalmente Humanidades y Ciencias Sociales (en todos los casos p-valor <0.05), no se encontraron diferencias entre Ciencias Biosanitarias y Humanidades (p>0.05).

En relación a la variable “descargas” se observó que existían diferencias estadísticamente significativas entre las áreas de Ciencias Biosanitarias y Ciencias Sociales y entre Ciencias Experimentales y Ciencias Sociales (en todos los casos p-valor <0.05), en el resto de las áreas no se encontraron diferencias. Por último, la variable “citación” presentó diferencias significativas en las áreas de Enseñanzas Técnicas y Ciencias Sociales, entre Ciencias Biosanitarias y Humanidades y entre Ciencias Biosanitarias y Ciencias Sociales (en todos los casos p-valor <0.05), no se encontraron diferencias en las otras áreas (véase Tabla 38).

Tabla 38

Características de los indicadores de visibilidad e impacto según las áreas

ÁREAS			
Total GLOBAL (n=863)	Media±DS	Mediana (RI)	p-valor
Visitas *&+#	263.23±195.60	228 (170-308)	<0.001
Descargas & +	2114.44±3661.12	1349 (814-2289)	<0.001
Citas *@&	0.60±2.01	0.00 (0.00-0.00)	<0.001
BIOSANITARIAS (n=249)			
Visitas	231.81±119.82	208 (160.5-280)	
Descargas	1780.35±1949.94	1228 (809-1903)	
Citas	0.27±1.90	0 (0-0)	
TÉCNICAS (n=32)			
Visitas	263.25±342.66	196 (163-248.25)	
Descargas	1814.47±2923.73	1214.5 (44.75-2036)	
Citas	0.13±0.55	0 (0-0)	
SOCIALES (n=309)			
Visitas	301.25±204.58	256 (195.5-352)	
Descargas	2696.78±5445.65	1492 (1008-2502.5)	
Citas	0.13±0.55	0.00 (0.00-1.00)	
HUMANIDADES (n=156)			
Visitas	258.21±252.30	222.5 (152.5-304.75)	
Descargas	1968.74±2006.54	1453 (581.5-2802)	
Citas	0.69±1.72	0.00 (0.00-0.75)	
EXPERIMENTALES (n=117)			
Visitas	234.25±138.10	219 (171.5-271)	
Descargas	1563.82±1889.30	1015 (568-1828.5)	
Citas	0.68±1.74	0 (0-0)	

Post-hoc: *: diferencias entre E. Técnicas y CC. Sociales. &: diferencias entre CC. Biosanitarias y CC. Sociales. +: diferencias entre CC. Experimentales y CC. Sociales. #: diferencias entre Humanidades y CC. Sociales. @: diferencias entre CC. Biosanitarias y Humanidades

Nota: Elaboración propia

En relación a los años, en la Tabla 39 se muestran los valores medios para los indicadores estudiados, observándose que hay diferencias estadísticamente significativas solo para la variable 'descargas' entre los años 2008 y 2009, 2008 y 2010, y 2008 y 2011, y entre 2009 y 2011 (p-valor <0,05), no presentando diferencias estadísticamente significativas el resto de las variables y años analizados (véase Tabla 39).

Tabla 39

Características de los indicadores de visibilidad e impacto según los años

Total GLOBAL (n=863)	AÑOS		
	Media±DS	Mediana (RI)	p-valor
Visitas	263.23±195.60	228 (170-308)	0.497
Descargas &*+##	2114.44±3661.12	1349 (814-2289)	<0.001
Citas	0.60±2.01	0.00 (0.00-0.00)	0.077
2008 (n=152)			
Visitas	260.24±160.43	223.5 (172.25-308)	
Descargas	3207.64±6311.11	1891.5 (886.5-3738.75)	
Citas	0.99±2.59	0.00 (0.00-1.00)	
2009 (n=155)			
Visitas	248.75±211.23	214 (163-296)	
Descargas	1760.45±1577.10	1365 (1011-2207)	
Citas	0.58±1.71	0 (0-0)	
2010 (n=297)			
Visitas	264.91±201.41	231 (173.5-307.5)	
Descargas	1801.25±2402.06	1348 (824-2008.5)	
Citas	0.57±2.37	0.00 (0.00-1.00)	
2011 (n=259)			
Visitas	271.73±198.44	236 (171-322)	
Descargas	2043.87±3531.09	1104 (675-2060)	
Citas	0.43±1.12	0.00 (0.00-0.00)	

Post-hoc: &: diferencias en la variable descargas para los años 2008 y 2009. *: diferencias en la variable descargas para los años 2008 y 2011. +: diferencias en la variable descargas para los años 2008 y 2011. #: diferencias en la variable descargas par los años 2009 y 2011

Nota: Elaboración propia

En relación al grado de asociación entre las variables analizadas, se observa que existe una correlación significativa y positiva, aunque no muy fuerte, entre las variables visitas, descargas y citas (a medida que aumentan las visitas, aumentan las descargas y las citas). En la Tabla 40 se indican las correlaciones establecidas entre las variables.

Tabla 40

Correlaciones entre las variables

	Visitas	Descargas	Citas	p-valor
Visitas	1	0.493**	0.256**	<0.01
Descargas	0.493**	1	0.237**	<0.01
Citas	0.256**	0.237**	1	<0.01

Nota: Elaboración propia

El porcentaje de tesis doctorales que han tenido alguna cita es del 21,7% frente al 78,3% de las que no han tenido ninguna (véase la Figura 43).

■ Al menos una cita ■ Ninguna cita

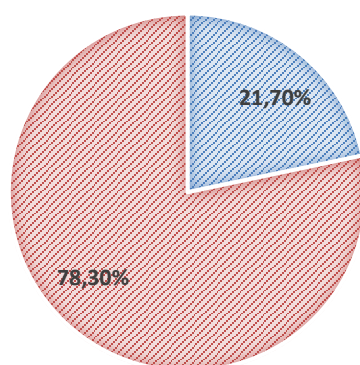


Figura 43. Porcentaje de citas. Fuente: Elaboración propia

A continuación, se muestra la distribución de las citas por Áreas, observándose cómo quedan distribuidas, así, por ejemplo, el 88,8 % de las tesis en OA del área biosanitaria no tiene ninguna cita y el 32,7% de las tesis que no tienen ninguna cita pertenecen a esta área.

De la misma manera se observa que el 30,1% de las tesis doctorales del área de Ciencias Sociales tienen al menos una cita y el 49,7% de las tesis que tienen al menos una cita pertenecen al área de Ciencias Sociales. En la Tabla 41 se muestra la distribución de las citas por áreas de conocimiento.

Tabla 41

Distribución de las citas por áreas de conocimiento

	Ninguna cita (n=676)	Al menos una cita (n=187)	p-valor
CC. Biosanitarias	221	28	<0.001
%dentro del área	88.8	11.2	
%Citas	32.7	15.0	
E. Técnicas	30	2	
%dentro del área	93.8	6.2	
%Citas	4.4	1.1	
CC. Sociales	216	93	
%dentro del área	69.9	30.1	
%Citas	32.0	49.7	
Humanidades	117	39	
%dentro del área	75.0	25.0	
%Citas	17.3	20.9	
CC. Experimentales	92	25	
%dentro del área	78.6	21.4	
%Citas	13.6	13.4	

Nota: Elaboración propia

A continuación, se muestra la distribución de las citas según años de defensa (Tabla 42). Se observa, por ejemplo, como el 79,8 % de las tesis doctorales en OA del año 2010 no tiene ninguna cita y el 35,1% de las tesis que no tienen ninguna cita fueron leídas en ese año.

De igual manera se observa que el 28,9% de las tesis doctorales leídas en el año 2008 tiene al menos una cita, y el 23,5% de las tesis que tienen al menos una cita fueron leídas ese año.

Tabla 42

Distribución de citas según años

	Ninguna cita (n=676)	Al menos una cita (n=187)	p-valor
2008	108	44	0.124
%dentro del año	71.1	28.9	
%Citas	16.0	23.5	
2009	124	31	
%dentro del año	80.0	20.0	
%Citas	18.3	16.6	
2010	237	60	
%dentro del año	79.8	20.2	
%Citas	35.1	32.1	
2011	207	52	
%dentro del año	79.9	20.1	
%Citas	30.6	27.8	

Nota: Elaboración propia

III.2.1. Cuestiones previas

La discusión se ha centrado en los resultados obtenidos en el estudio realizado a través de los distintos sistemas de análisis y herramientas utilizados para ello. Se destacan solo aquellos resultados que han sido significativos o que han aportado alguna novedad al tema tratado de la visibilidad, el uso y el impacto de las tesis doctorales en acceso abierto.

El punto de partida de esta investigación fue estudiar los aspectos de visibilidad, uso e impacto de las tesis doctorales leídas en la Universidad de Salamanca entre los años 2006 y 2011.

Este período de tiempo comprende lo que se podría denominar fase pre-repositorio o No OA, años 2006 y 2007, en la que no existía el compromiso institucional con el acceso abierto y, por tanto, las tesis presentadas en estos años no estaban sujetas a mandato institucional y no se difundían en abierto; y, por otro lado, la fase OA que comprende los años 2008-2011 donde estaba vigente el mandato institucional por el que se requería que todas las tesis presentadas en la Universidad de Salamanca estuvieran depositadas en el repositorio, en acceso abierto. A partir de 2012 la normativa cambia con la entrada en vigor del Real Decreto 99/2011 y es este el motivo por el que las tesis leídas a partir de este año no se han recogido en la muestra.

De acuerdo con Sale (2006), aunque los mandatos tardan tiempo en ser integrados en los procesos de trabajo de los académicos y en las instituciones, se demuestra en este trabajo de investigación doctoral que son una forma efectiva para lograr el crecimiento y garantizar la sostenibilidad de las colecciones depositadas en los repositorios.

Asimismo, se considera que el mandato no daña la imagen del repositorio como servicio integral a la comunidad académica, en contra de la opinión de Palmer, Tefreau y Newton (2008), autores que apuntaban la idea de que es más importante la visión del personal académico que la

garantía de asegurar el regular crecimiento de contenido del repositorio. Por el contrario, el mandato no solo asegura el crecimiento del contenido del repositorio como aportación obligada de la comunidad investigadora al bien público, sino que también asegura para los autores las ventajas derivadas de colaborar y compartir, tal como se ha demostrado con el análisis de visibilidad, uso e impacto llevado a cabo en esta tesis doctoral.

La elección de este tipo de literatura gris se debe a que es un tipo de contenido científico, alojado en la casi totalidad de los repositorios institucionales mundiales, sobre el que no hay demasiados estudios de visibilidad e impacto, al contrario de lo que sucede con los artículos científicos.

Otra razón para realizar este estudio se encontró en la obra *El acceso abierto* de Suber (2015) donde el autor defiende que “las tesis son el tipo más útil de la producción científica invisible y la más invisible de los tipos de producción útil y, que debido a su alta calidad y baja visibilidad, vale la pena resolver el problema para su acceso”.

Se consideró, por tanto, que era necesario un estudio bibliométrico que aportara datos que corroborarán que el acceso abierto es un modelo válido para la difusión de la literatura gris, concretamente para las tesis electrónicas alojadas en los repositorios institucionales porque aumenta la visibilidad, uso y citación en relación con las tesis que no están en acceso abierto.

III.2.2. WOS como herramienta de búsqueda de tesis citadas

En el desarrollo del estudio se comprobó que las tesis doctorales que no estaban en acceso abierto en el repositorio Gredos eran muy difíciles de rastrear, prácticamente imposible, ante la inexistencia de datos de uso sobre ellas.

La única forma de obtener datos de uso de las tesis doctorales que no están publicadas y no están en el repositorio podría ser comprobar el registro de consultas *in situ* del archivo universitario donde se depositan las tesis, y posiblemente las peticiones de consulta y préstamo

recibidas en el Servicio de Préstamo Interbibliotecario de la Universidad. Estas comprobaciones no se llevaron a cabo ante la falta de automatización de los procesos que permitieran recuperar de forma rápida y fiable los datos.

En cuanto al rastreo de las citas de las tesis doctorales No-OA se siguió un trabajo de Larivière, Zuccala y Archambault (2007), autores que utilizaron WOS como herramienta para localizar las tesis citadas de una determinada universidad en un período de tiempo determinado.

En la búsqueda realizada en esta base de datos, solo se recuperaron 66 tesis doctorales que tenían alguna cita en 77 trabajos (Figura 37, Índice de referencias citadas), cifra muy poco representativa considerando que el total de tesis doctorales de la Universidad de Salamanca defendidas en el período de tiempo analizado es de 1.435 tesis. Además, los datos recuperados en WOS de las tesis doctorales son bastante incompletos y en algunos casos hacen imposible la identificación de los trabajos. En 27 casos fue imposible la identificación de las tesis doctorales porque los registros resultantes carecían de título y la identificación por autor también se hacía imposible.

A través de estos resultados, se observa que la mayor parte de las citas recaen sobre tesis doctorales que están publicadas o que están en acceso abierto, es decir que son fácilmente recuperables y consultables a través de Internet. De todas formas, las tesis doctorales tienen muy poca citación en las publicaciones científicas como indican Larivière, Zuccala y Archambault (2008) y como se ha demostrado con esta búsqueda. La razón de este hecho puede deberse a las preferencias de los autores por citar artículos y libros derivados de la investigación, en lugar de citar las tesis doctorales originales.

Asimismo, el hecho de que cada vez haya más e-tesis doctorales disponibles en acceso abierto requiere que se mida el uso de las mismas con herramientas adecuadas que proporcionen datos reales sobre visibilidad y citación. Según algunos autores las estadísticas indican que las tesis

doctorales que están disponibles en repositorios de acceso abierto son las más utilizadas y citadas en la investigación en curso (Stanton & Liew, 2011; Troman et al., 2007).

Finalmente, se observó a través de este análisis que las bases de datos WOS y Scopus no son herramientas adecuadas para facilitar datos sobre citación de este tipo de literatura gris científica esté o no disponible en acceso abierto.

III.2.3. Medición de la visibilidad, el uso y el impacto de las tesis doctorales en acceso abierto

Además de los datos sobre visibilidad, uso e impacto, del análisis de los resultados del estudio se extraen otros datos significativos sobre el comportamiento de la comunidad universitaria con respecto al acceso abierto y particularmente con respecto a las tesis doctorales depositadas en el repositorio.

Depósito de tesis doctorales en el repositorio

Como ya se ha visto, la muestra analizada está compuesta por 863 tesis doctorales leídas en la Universidad de Salamanca en el periodo comprendido entre los años 2008 y 2011. Si bien, de estas se analizaron finalmente 815, es decir, un 94,44% de la población inicial, ya que 48 tesis doctorales habían sido retiradas del repositorio a petición de los autores en diversos periodos de tiempo. El 5% de los autores retiraron su tesis doctoral del repositorio, alegando en todos los casos que existía un compromiso con editores para publicar y estos retenían los derechos de explotación por lo que no se permitía la difusión en acceso abierto.

El alto porcentaje de depósito de tesis doctorales en el repositorio se debe fundamentalmente al mandato existente desde 2008 en la Universidad de Salamanca, corroborando así la afirmación de muchos autores y de la doctoranda que defienden que las políticas de mandato contribuyen al aumento de contenidos en acceso abierto y, por ende, al aumento de visibilidad, uso y aumento

de citación sobre los autores y las instituciones (Anne Kennan, 2011; Anne Kennan & Wilson, 2006; É. Archambault et al., 2014; Gargouri, Lariviere, et al., 2012; Melero, 2013; Merlo-Vega & Ferreras-Fernández, 2010; Pinfield, 2015; Xia et al., 2012) También se observa, a través de ese 5% de tesis retiradas del repositorio, que existe un escaso número de autores que aún desconfían del acceso abierto o que desconocen sus derechos como autores alegando, como motivo para no poner en abierto su investigación doctoral, que los editores retienen los derechos de explotación sobre su tesis doctoral porque va a ser publicada.

Se ha comprobado que las áreas con mayor número de tesis doctorales leídas entre 2008-2011 fueron las áreas de Ciencias Sociales y de Ciencias Biosanitarias, datos que indican que estas áreas tuvieron mayor volumen de investigación doctoral que el resto. Por el contrario, destaca el área de Enseñanzas técnicas por el escaso número de tesis doctorales presentadas (32) en comparación con el resto de las áreas.

Visibilidad y uso

La visibilidad de las tesis doctorales del repositorio Gredos está estrechamente unida a la interoperabilidad alcanzada por el repositorio. Como afirma Hanief Bhat (2010) una característica primordial de los repositorios es que sean interoperables y el grado de visibilidad del contenido de los repositorios está estrechamente relacionado con la interoperabilidad de los mismos. A lo largo de esta investigación doctoral se ha tratado el tema de la interoperabilidad de los repositorios relacionada con el aumento de visibilidad, uso y citación, y también se ha apuntado a la posibilidad, no muy lejana, de que los repositorios lleguen a ser invisibles para el usuario y pasen a actuar en un segundo plano como apoyo a los proveedores de servicios externos y a facilitar la visibilidad del contenido a través de agregadores, indexadores y robots de búsqueda. Esta es la visión que subyace en el concepto “ecosistema tecnológico” y en el que el repositorio evoluciona e interacciona con usuarios y otros componentes, consiguiendo la interoperabilidad

mediantes servicios y definiciones ontológicas de los flujos de datos (García-Holgado & García-Peñalvo, 2013a, 2013b, 2016).

Google Analytics también proporciona datos que evidencian la visibilidad y el descubrimiento del repositorio gracias a la interoperabilidad (Cheng, Cheung, & Kot, 2014). Esta herramienta muestra las fuentes desde las que se accede al repositorio Gredos, poniendo de manifiesto que la interoperabilidad del repositorio con otros sitios contribuye a incrementar las visitas: Europea, Hispana, *Google Scholar*, Tesis en Red (TDR), DART-Europe, BASE, OATD, y otros.

En el estudio realizado, la visibilidad y el uso de las tesis doctorales de cada una de las cinco áreas de conocimiento, en las que estas fueron agrupadas, se midieron a través del número de visitas y de descargas recibidas por cada uno de los ítems de esta colección del repositorio. En general, reciben más visitas las tesis del área de Ciencias Sociales y destaca el aumento de la media de las visitas en el año 2010 en el área de las Enseñanzas técnicas, y en el año 2009 en el área de las Humanidades. Se ha observado que existían diferencias estadísticamente significativas entre las áreas de Enseñanzas Técnicas y Ciencias Sociales, Ciencias Biosanitarias y Ciencias Sociales, Ciencias Experimentales y Ciencias Sociales y finalmente Humanidades y Ciencias Sociales; mientras que no se encontraron diferencias entre Ciencias Biosanitarias y Humanidades.

En cuanto a la media de las descargas, es muy similar en todas las áreas de conocimiento, excepto en el área de Ciencias Sociales que en el año 2008 se produce un aumento de la media de las descargas recibidas. Al comparar las distintas áreas se observa que existen diferencias estadísticamente significativas entre las áreas de Ciencias Biosanitarias y Ciencias Sociales y entre Ciencias Experimentales y Ciencias Sociales, y que en el resto de las áreas no se encontraron diferencias.

Las tesis doctorales del área de la Ciencias Sociales y de las Humanidades son las más utilizadas, aunque existe muy poca diferencia con las otras áreas. Este hecho, podría deberse a que la literatura gris en acceso abierto, y particularmente las tesis doctorales, constituyen una importante fuente de información para la investigación en Ciencias Sociales y Humanidades.

A través de los datos de las estadísticas de visitas y descargas del repositorio Gredos es evidente un uso extenso de las tesis doctorales. En este punto haría falta preguntarse, al igual que Stone y Lowe (2014) si las tesis se utilizan y se citan en fuentes reconocidas y en revistas científicas, y por otro lado dónde acaba la investigación de los doctorandos.

Citación

Al igual que Vashney (2012) la doctoranda afirma que la información difundida a través de Internet y recuperable a través de *Google*, aumenta las posibilidades de ser citada, lo cual es una oportunidad para las tesis doctorales depositadas en repositorios de acceso abierto.

A través del estudio, se ha observado que en relación a la citación existían diferencias significativas entre las áreas de Enseñanzas Técnicas y Ciencias Sociales, entre Ciencias Biosanitarias y Humanidades y entre Ciencias Biosanitarias y Ciencias Sociales; en las otras áreas no se encontraron diferencias. En relación a los años recogidos en el análisis, se observó que no existen diferencias significativas en cuanto a visitas y citación, pero sí existen en cuanto a descargas entre los años 2008 y 2009, 2008 y 2010, y 2008 y 2011, y entre 2009 y 2011.

En relación al grado de asociación entre las variables analizadas, se observa que existe una correlación significativa y positiva, aunque no muy fuerte, entre las variables visitas, descargas y citas (a medida que aumentan las visitas, aumentan las descargas y las citas).

En cuanto al impacto de la literatura gris en la *Web*, Cesare y otros (2008) concluyen que el sistema de rastreo de las citas de *Google Scholar* da visibilidad a la literatura gris, tanto como

documento citado como que cita, confirmando su valor en la comunicación académica. Además *Google Scholar* también permite reconstruir el ciclo de vida del documento al hacer posible que se encuentren las diferentes versiones a texto completo del documento, haciendo así evidente la transformación de la literatura gris en literatura convencional.

III.2.4. *Google Scholar* como herramienta alométrica para la identificación de citas

Para esta investigación, *Google Scholar* (GS) ha sido una aplicación muy útil en la identificación de citas de las tesis doctorales del repositorio Gredos, que de otra forma hubiese sido imposible localizar. Existían ya algunos trabajos sobre la medición con *GS* que mostraban la utilidad de esta herramienta para medir el impacto de los trabajos en acceso abierto (C. Burns, 2013; Cesare et al., 2008; “EC3noticias: *Google Scholar*: presente y futuro, fortalezas y debilidades, reportaje de John Bohannon en *Science*,” n.d.; Falagas, Pitsouni, Malietzis, & Pappas, 2007; Galindo & Llopes, 2015; Jacsó, 2012; Kousha & Thelwall, 2007a, 2007b; Martin-Martin et al., 2016; Orduña-Malea & López-Cózar, 2014, 2016; Prins, Costas, van Leeuwen, & Wouters, 2016; Rehemtula, Rosa, Leitão, & Avilés, 2014).

Al igual que *Scopus*, *Google Scholar* (GS) también hizo su aparición en 2004. Harzing y Van der Wall (2007) describen a *Google Scholar* como el responsable de la democratización del análisis de citas. Generalmente, las fuentes de citación son usadas para realizar un seguimiento de lo que está haciendo alguien determinado en un campo de estudio específico y el impacto de la investigación dentro de la disciplina. El gran crecimiento de los recursos de citas en línea, como *Scopus* y *GS* han traído consigo nuevas oportunidades para los académicos en cuanto a las citas de seguimiento y al análisis de citas (Adriaanse & Rensleigh, 2013).

De acuerdo con Torres-Salinas, Ruiz-Pérez y Delgado-López-Cózar (2009), *Google Scholar* en primer lugar es un buscador de publicaciones científicas y, en segundo lugar, es un índice de

citas que ayuda a conocer el impacto que tienen las publicaciones. Esta última característica es la más interesante y la que lo convierte en un competidor directo de otros índices de citación como WOS, o Scopus. Por estas funciones GS se presenta como una aplicación ideal para buscar el texto completo de un trabajo, buscar la producción bibliográfica de un autor, de una revista o sobre un tema y, buscar las citas que recibe un trabajo (libro, artículo de revista, tesis, informe...).

Estos autores consideran que GS es útil sobre todo es para la literatura no anglosajona, que es la peor controlada por los sistemas de información dominantes en el mundo de la ciencia, para las disciplinas que no emplean preferentemente las revistas como medio de comunicación (ingenierías, humanidades, ciencias sociales...) y para localizar citas a libros, tesis, informes y a artículos publicados en revistas secundarias no incorporadas a la llamada “corriente principal de la ciencia” (Torres-Salinas et al., 2009).

Galloway y Pease (2013) consideran que *Google Scholar Citations* es una de las herramientas más maduras y prometedoras de las alométricas, destinada a rastrear el número de lectores y la influencia. Los autores la describen como un servicio que permite a los autores rastrear sus publicaciones e influencia usando *Google Scholar Metrics*. Se trata de una herramienta libre, extremadamente útil, de uso amigable y bien considerada.

Módulo IV. Conclusiones

En el módulo IV se presentan las conclusiones que se desprenden de la investigación realizada en esta tesis doctoral. En el capítulo 1 se recogen las conclusiones separadas y numeradas; por una parte se exponen, las conclusiones generales relacionadas directamente con los objetivos propuestos en esta tesis doctoral y, por otra parte, las conclusiones relacionadas con las preguntas de investigación planteadas. Seguidamente, se establece una conclusión final que engloba y resume todas las anteriores.

Asimismo, en el capítulo 2 de este módulo se recogen las publicaciones de la doctoranda relacionadas con la investigación realizada. La práctica de incluir las publicaciones en el módulo de conclusiones, es una recomendación del Programa de Doctorado *Formación en la Sociedad del Conocimiento*, en el que se inscribe esta tesis doctoral.

Finalmente, en el capítulo 3 se indican las investigaciones futuras que se llevarán a cabo a partir de la finalización de esta tesis.

Capítulo IV.1. Conclusiones

Conclusiones generales

1. Se han alcanzado los objetivos y sub-objetivos planteados en esta investigación:

- Se ha definido el estado de la cuestión de los repositorios institucionales de acceso abierto tanto desde un punto de vista estratégico-político como tecnológico.
- Se ha definido el estado de la cuestión de las tesis doctorales electrónicas en acceso abierto.
- Se ha demostrado que la interoperabilidad de los repositorios (aspecto técnico) y las políticas de mandato afectan positivamente a la visibilidad, el uso y el impacto de la investigación depositada en los repositorios institucionales y por ende a las tesis electrónicas
- Se ha realizado un estudio cuantitativo en el repositorio Gredos que tome como elemento de estudio las tesis doctorales de la Universidad de Salamanca.
- Se han estudiado y aplicado métricas alternativas para la realización del estudio.
- Se ha establecido un modelo bibliométrico viable para la medición del uso y de la citación de las tesis depositadas en acceso abierto.

2. El acceso abierto reporta beneficios de visibilidad e impacto a las tesis doctorales. Al comienzo de esta tesis doctoral se planteaba la cuestión sobre los beneficios del acceso abierto para las tesis doctorales depositadas en repositorios institucionales. Los beneficios de visibilidad e impacto han sido suficientemente demostrados a través de los resultados obtenidos en el análisis. Las tesis en acceso abierto obtienen altas tasas de visibilidad y uso (visitas y descargas) y también obtienen una pequeña tasa de citación. Las tesis que se difunden a través de los repositorios se benefician de la interoperabilidad de estos, porque facilitan la difusión a

través de múltiples portales, sitios y robots de búsqueda aumentando de esta forma su visibilidad, uso e impacto.

3. Los resultados demuestran que los repositorios son una vía válida y ventajosa para difundir la literatura gris científica de las instituciones. Los repositorios proporcionan información sobre uso (visibilidad) y sobre citación (impacto) de las tesis doctorales, esta información no es posible en el caso de las tesis que no están en acceso abierto.

4. El uso y el impacto de las tesis doctorales en repositorios significa que las tesis doctorales ya no son literatura gris, sino que se trata de literatura visible, consultada y citada que contribuye a la visibilidad de autores e instituciones.

5. A menudo, las tesis doctorales se consideran información de segundo nivel cuando se comparan con los artículos publicados en revistas científicas. Esto se debe a dos razones fundamentalmente: la primera es que las tesis no tienen revisión por pares y la segunda que se realizan al principio de una carrera académica. Sin embargo, se ha demostrado que las tesis doctorales no solo producen resultados que reflejan años de investigación intensa combinada con una rica revisión de la literatura, sino que también representan una parte cada vez mayor de contenido abierto disponible en los repositorios institucionales, lo que contribuye al impacto global y al *ranking* de la institución.

Conclusiones relacionadas con las preguntas de investigación

1. Las tesis doctorales puestas en acceso abierto son más visibles y más usadas que las tesis doctorales que no son de acceso abierto, como se ha comprobado a través de los resultados obtenidos en el análisis de visibilidad y uso de esta tesis doctoral.

Las tesis doctorales que no están publicadas ni disponibles en abierto no son visibles, como se ha comprobado a través de las búsquedas realizadas en la base de datos WOS. No existen datos que permitan analizar el uso de estas tesis.

2. Las tesis doctorales en acceso abierto son más citadas que las tesis que no están en acceso abierto. Las tesis doctorales que no están publicadas y que no se difunden en acceso abierto no se ven, no se utilizan y por tanto no se citan.

Los resultados del análisis de citación también han demostrado que las tesis doctorales en acceso abierto tienen citación, aunque la relación no sea proporcional con los datos de uso (visitas y descargas).

3. Interoperabilidad y mandato son dos factores que influyen positivamente en la visibilidad, el uso y el impacto de las tesis doctorales.

Por un lado, la interoperabilidad de los repositorios (aspecto técnico) influye en el aumento de visibilidad, el uso y el impacto de las tesis doctorales depositadas en los repositorios. La interoperabilidad del repositorio Gredos ha facilitado la difusión de los contenidos de la colección de tesis doctorales a través de recolectores y buscadores científicos, aumentando de esta forma el número de visitas y de descargas (visibilidad y uso) y posibilitando la citación.

Por otro lado, al igual que sucede con otros tipos de trabajos de investigación publicados y puestos en acceso abierto, las políticas de mandato (aspecto político) que obligan a depositar las tesis doctorales en el repositorio institucional de acceso abierto afectan positivamente al aumento de contenidos en el repositorio y por tanto al aumento de su visibilidad, uso e impacto. El alto porcentaje de depósito de tesis doctorales en el repositorio Gredos se debe fundamentalmente al mandato de la Universidad de Salamanca, corroborando así la afirmación de que las políticas de mandato contribuyen al aumento de contenidos en acceso abierto y, por ende, al aumento de visibilidad, uso y aumento de citación sobre los autores y las instituciones

4. La visibilidad, usabilidad e impacto del acceso abierto se mide con herramientas tradicionales, tales como los módulos de estadísticas de los repositorios para medir el uso, y con métricas alternativas, tales como *Google Scholar Citations*.

El módulo de estadísticas de Gredos ha permitido extraer datos de visitas y descargas para analizar el uso de las tesis doctorales en abierto. *Google Scholar* ha permitido obtener datos sobre el número de citas recibidas para analizar el impacto de las tesis doctorales.

5. No se puede realizar un análisis comparativo de impacto de las tesis en acceso abierto con las tesis que no están en acceso abierto, dada la imposibilidad de medir el uso y citación de las tesis doctorales que no están en acceso abierto.

A través del análisis realizado se ha comprobado que no es posible aplicar herramientas de medición de uso e impacto a las tesis doctorales que no están en acceso abierto, por tanto no es viable realizar un estudio comparativo con las tesis que se difunden en abierto.

Conclusión final

Las tesis doctorales son trabajos de investigación que contienen resultados y una rica revisión de la literatura, y como se ha observado a través del estudio de caso del repositorio Gredos representan una parte muy importante del contenido de este repositorio, contribuyendo al impacto global y al *ranking* de la Universidad de Salamanca. Así pues, los repositorios institucionales son un canal válido para la difusión de la literatura gris científica, particularmente de las tesis doctorales.

El sistema de análisis bibliométrico utilizado en esta investigación es viable para la medición de la visibilidad, el uso, la citación y el impacto de las tesis depositadas en repositorios institucionales de acceso abierto. Las tesis doctorales son citadas aunque la relación entre uso y citas recibidas no es proporcional, hecho que permite deducir la falta de buenas prácticas de citación sobre la literatura gris.

Capítulo IV.2. Publicaciones

En este apartado se indican las principales publicaciones relacionadas con la investigación desarrollada en esta tesis doctoral.

Artículos en revistas científicas

Ferreras Fernández, T. (2013). El Repositorio Gredos de la Universidad de Salamanca como proveedor de datos de Europeana. *Boletín de la ANABAD*, 63(3), 407-419.

Ferreras-Fernández, T., García-Peñalvo, F. J., Merlo-Vega, J.A., & Martín-Rodero, H. (accept, 21-Jul-2016). Providing Open Access to PhD theses: Visibility and citation benefits. *Program* 50(4).

Ferreras-Fernández, T., & Merlo-Vega, J. A. (2015). Repositorios de acceso abierto: un nuevo modelo de comunicación científica. *La Revista de la Sociedad ORL CLCR en el repositorio Gredos. Rev. Soc. Otorrinolaringol. Castilla León Cantab. La Rioja*, 6(12), 94 -113.

García-Peñalvo, F. J., Merlo-Vega, J. A., Ferreras-Fernández, T., Casaus-Peña, A., Albás-Aso, L., & Atienza-Díaz, M. L. (2010). Qualified Dublin Core Metadata Best Practices for GREDOS. *Journal of Library Metadata*, 10(1), 13–36. <http://doi.org/10.1080/19386380903546976>.

Merlo-Vega, J.A., & Ferreras-Fernández, T. (2013). Preservación digital y difusión de la revista Educación y biblioteca desde el repositorio Gredos de la Universidad de Salamanca. *El profesional de la información*, 22 (2), 143-148.

Ríos-Hilario, A., Martín-Campo, D., & Ferreras-Fernández, T. (2012). Linked data y linked open data: su implantación en una biblioteca digital. El caso de Europeana. *El profesional de la información*, 21 (3), 292-297.

Ríos-Hilario, A., Ferreras-Fernández, T., & Martín-Campo, D. (2014). From Bibliographic Records to Data Changes in the Library Environment with the Application of Linked Open Data Technologies. *Information Resources Management Journal*, 27(3), 28-41. DOI: 10.4018/irmj.2014070103.

Morales-Morgado, E. M., Campos Ortuño, R. A., Ling, Y. Ling, & Ferreras-Fernández, T. (2014). Adaptation of Descriptive Metadata for Managing Educational Resources in the GREDOS Repository. *International Journal of Knowledge Management (IJKM)*, 10(4), 50-72. DOI:10.4018/ijkm.2014100104

Capítulos de libros

Ferreras Fernández, T. (2012). Libros electrónicos en acceso abierto. En Cordón-García, José Antonio, Fernando Carbajo Cascón, Alonso-Arévalo, Julio. (coords.), *Libros electrónicos y contenidos digitales en la sociedad del conocimiento: mercado, servicios y derechos*, Madrid, Pirámide, 2012. ISBN 978-84-368-2769-9.

Tesis de Máster

Ferreras Fernández, Tránsito (2010). *Preservación digital en repositorios institucionales: GREDOS*. Tesis de Máster. Disponible en <http://hdl.handle.net/10366/83130>

Editoriales, informes, y otras contribuciones

Ferreras Fernández, T. (2011). Editorial "Acceso abierto al conocimiento, políticas gubernamentales y derechos de autor" *Scopeo*, n.41. Disponible en <http://scopeo.usal.es/boletin/41>

Merlo-Vega, J. A., & Ferreras-Fernández, T. (2016). Colaboración Bibliotecas de la USAL y Sociedad Otorrinolaringológica de Castilla y León, Cantabria y La Rioja. *Revista ORL*, 7(1). presentado en 04/2016. doi:<http://dx.doi.org/10.14201/orl201671.13875>

Ferreras Fernández, T. (et al.) Ciencia 2.0: aplicación de la web social a la investigación, <http://hdl.handle.net/10366/102598>

Ferreras Fernández, T. (2012) Dublin Core Cualificado: documento de trabajo. Ciencias de la Información, 43(1). <http://hdl.handle.net/10366/56268>

Ferreras-Fernández, T. (2013). La interoperabilidad: el pegamento técnico para unir repositorios. Blok de Bid, 27/03/2013. Disponible en: <http://hdl.handle.net/10366/122133>.

Actas de Congresos

Ferreras-Fernández, T., Merlo-Vega, J.A., & García-Peñalvo, F.J. (2013). Impact of scientific content in open access institutional repositories: a case study of the repository Gredos. In Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM '13), Francisco José García-Peñalvo (Ed.). ACM, New York, NY, USA, 357-363. DOI=10.1145/2536536.2536590

Ferreras-Fernández, T., Merlo-Vega, J.A. & García-Peñalvo, F.J. (2013). Science 2.0 supported by open access repositories and open linked data. In Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM '13), Francisco José García-Peñalvo (Ed.). ACM, New York, NY, USA, 331-332. DOI=10.1145/2536536.2536586
Disponible en: <http://doi.acm.org/10.1145/2536536.2536586>

Merlo Vega, J.A., y Ferreras Fernández, T. (2013). Cooperación para la preservación digital de contenidos locales: experiencias de las bibliotecas de las universidades públicas de Castilla y León. En VI Congreso Nacional de Bibliotecas Públicas, celebrado en Burgos 9-11 de octubre de 2012 (pp. 445-450). Madrid: Ministerio de Cultura, Educación y Deporte, 2013. NIPO 030-13-016-0.

Morales-Morgado, E. M., Campos Ortuño, R. A., Ling, Y. Ling, & Ferreras-Fernández, T. (2013). Metadata mapping to describe learning objects and educational apps in the GREDOS repository. In Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM '13), Francisco José García-Peñalvo (Ed.). ACM, New York, NY, USA, 349-356. DOI=10.1145/2536536.2536589

Ríos-Hilario, A., Ferreras-Fernández, T., & Martín-Campo, D. (2013). Linked open bibliographic data. In Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM '13), Francisco José García-Peñalvo (Ed.). ACM, New York, NY, USA, 333-337. DOI=10.1145/2536536.2536587

Capítulo IV.3. Investigaciones futuras

A través del trabajo llevado a cabo en esta tesis doctoral se ha comprobado que algunos aspectos de la investigación quedaban incompletos porque no se podían ampliar, o simplemente no era pertinente acometerlos por no estar dentro de los objetivos marcados en la investigación. Por ello la doctoranda se plantea seguir trabajando en futuras investigaciones sobre tres temas o tópicos fundamentalmente y que se exponen a continuación.

1. Revisión sistemática de la literatura. Desarrollar una metodología para la elaboración de revisiones sistemáticas en el ámbito de la Información y Documentación, ya que se ha comprobado que el empleo de las revisiones sistemáticas en esta área no está aún muy extendido.

2. La literatura gris de la Universidad de Salamanca como contribuyente al trabajo científico. Analizar el uso de la literatura gris de la Universidad de Salamanca en las referencias bibliográficas de los artículos de revistas con el fin de verificar en qué medida la literatura gris contribuye al trabajo científico. Siguiendo la línea de trabajo de Cesare y Ruggieri (2006) que aplicaron una metodología de análisis de citación de literatura gris para descubrir su impacto en diferentes disciplinas a través del uso de indicadores bibliométricos similares a los utilizados para la literatura convencional, tales como la frecuencia de uso y la frecuencia de cita de la literatura gris.

También es de interés profundizar en las prácticas de citación de los autores respecto a las tesis doctorales. Investigar las causas de la baja tasa de citación en comparación con las altas tasas de uso (visitas y descargas).

3. Análisis del impacto de los trabajos derivados de las tesis doctorales. El análisis llevado a cabo en esta investigación se ha centrado exclusivamente en las tesis doctorales difundidas a través de un RI. Se ha demostrado que es imposible determinar el impacto de las tesis doctorales que no están en acceso abierto, ya que se citan a menudo a través de documentos que son el resultado de las tesis en lugar de la tesis en sí. Se ha demostrado también que es imposible realizar una comparación de visibilidad e impacto entre las tesis doctorales OA y las tesis doctorales que no son OA porque los sistemas de medida que se han utilizado en esta investigación solo son válidos para las tesis doctorales en acceso abierto. Se ha visto que la metodología utilizada por Larivière, Zuccala y Archambault (2008) no era efectiva para medir las tesis que no están publicadas ni en acceso abierto. La metodología apropiada para obtener datos de citación de las tesis que no están en acceso abierto y también de las que lo están podría ser la de rastrear, localizar y analizar los “trabajos derivados de las tesis”, como son los artículos, capítulos de libros y conferencias.

Para la identificación de los trabajos derivados se podría utilizar la encuesta a los autores. Sin embargo, este método puede resultar bastante engorroso y poco fiable debido al gran tamaño de la muestra, por tanto parece más adecuada la técnica empleada por Echeverria, Stuart y Blanke (2015) con el fin de detectar artículos derivados mediante la identificación de ítems a través del análisis basado en el texto completo de las tesis doctorales y de los artículos con los que comparten autoría. Después de identificar las obras publicadas derivadas de las tesis, habría que realizar un seguimiento de citación en las bases de datos e índices para analizar su impacto.



MÓDULO V. BIBLIOGRAFÍA

MÓDULO V. Bibliografía

- Abadal, E. (2012). *Acceso abierto a la ciencia*. Barcelona: UOC.
- Abadal, E., Ollé Castellà, C., Abad-García, F., & Melero, R. (2013). Políticas de acceso abierto a la ciencia en las universidades españolas. *Revista española de Documentación Científica*, 36(2), e007. <http://doi.org/10.3989/redc.2013.2.933>
- Adriaanse, L. S., & Rensleigh, C. (2013). Web of Science, Scopus and Google Scholar. *The Electronic Library*, 31(6), 727–744. <http://doi.org/10.1108/EL-12-2011-0174>
- Agudelo, D., Ortiz-Recio, G., Fullerat, C. V., Valor-Segura, I., Bretón-López, J., Poveda-Vera, J., & Álvarez, I. T. (2003). Análisis de la productividad científica de la Psicología española a través de las tesis doctorales. *Psicothema*, 15(4), 595–609.
- Aguillo, I. F., Ortega, J. L., Fernández, M., & Utrilla, A. M. (2010). Indicators for a webometric ranking of open access repositories. *Scientometrics*, 82(3), 477–486.
- Alier Forment, M., Casan Guerrero, M. J., Conde González, M. A., García-Peñalvo, F. J., & Severance, C. (2010). Interoperability for LMS: the missing piece to become the common place for e-learning innovation. *International Journal of Knowledge and Learning*, 6(2/3), 130. <http://doi.org/10.1504/IJKL.2010.034749>
- Alier Forment, M., Mayol, E., Casany, M. J., Piguillem, J., Merriman, J. W., Conde González, M. Á., ... Severance, C. (2012). Clustering Projects for eLearning Interoperability. *JUCS - Journal of Universal Computer Science*, 18(1), 106–122. <http://doi.org/10.3217/jucs-018-01-0106>
- American Psychological Association. (2010). *Publication Manual of the American Psychological Association*. (American Psychological Association, Ed.) (6th ed.).
- Andersson, S., & Svensson, A. (2013). *Repositories Recreated: Working Towards Improved Interoperability and Integration by a Co-operative Approach in Sweden*. In ELPUB 2013, 17th International Conference on Electronic Publishing, Mining the Digital Information Networks June 13-14, 2013, Blekinge Institute of Technology, Karlskrona, Sweden.
- Anne Kennan, M. (2011). Learning to share: mandates and open access. *Library Management*, 32(4/5), 302–318. <http://doi.org/10.1108/01435121111132301>
- Antelman, K. (2004). Do Open-Access Articles Have a Greater Research Impact? *College & Research Libraries*, 65(5), 372–382. <http://doi.org/10.5860/crl.65.5.372>
- Archambault, E., Amyot, D., Deschamps, P., Nicolle, A., & Roberge, G. (2013). *Proportion of Open Access Peer-Reviewed Papers at the European and World Levels—2004-2011*. Montreal.
- Archambault, É., Caruso, J., & Nicol, A. (2014). *State-of-art analysis of OA strategies to peer-review publications*. Montreal.
- Arguimbau Vivó, L., & Fuentes Pujol, M. E. (2009). Tesis doctorales en España: análisis de la bibliografía específica y su accesibilidad. *Ibersid*, 95–104.

Armbruster, C., & Romary, L. (2010). Comparing Repository Types - Challenges and barriers for subject-based repositories, research repositories, national repository systems and institutional repositories in serving scholarly communication. *International Journal of Digital Library Systems*, 1(4), 61–73.

Association of College & Research Libraries. (2003). Principles and Strategies for the Reform of Scholarly Communication 1. Retrieved from <http://www.ala.org/acrl/publications/whitepapers/principlesstrategies>

Atchison, A., & Bull, J. (2015). Will Open Access Get Me Cited? An Analysis of the Efficacy of Open Access Publishing in Political Science. *PS: Political Science & Politics*. Retrieved from http://scholar.valpo.edu/ccls_fac_pub/21

Auger, C. P. (Ed.). (1975). *Use of Reports Literature*. London: Butterworth.

Bar-Ilan, J. (2008). Informetrics at the beginning of the 21st century—A review. *Journal of Informetrics*, 2(1), 1–52. article. <http://doi.org/10.1016/j.joi.2007.11.001>

Barrueco Cruz, J. M. (2008). Medición del uso e impacto de documentos distribuidos a través de repositorios institucionales. *BiD: Textos Universitaris de Biblioteconomia I Documentació*, (20).

Barrueco Cruz, J. M., & García Testal, C. (2009). Repositorios institucionales universitarios: evolución y perspectivas. *Zaragoza: Fesabid, XI Jornadas*.

Baruch, P. (2007). La diffusion libre du savoir Accès libre et Archives ouvertes. *L'Archicube*, 3, 77–95.

Baruch, Y., Ghobadian, A., & Ozbilgin, M. (2013). Open Access the Wrong Response to a Complex Question: The Case of the Finch Report. *British Journal of Management*, 24(2), 147–155. <http://doi.org/10.1111/1467-8551.12016>

Bell, S., Foster, N., & Gibbons, S. (2005). Reference Librarians and the Success of Institutional Repositories. *Reference Service Review*, 33(3), 283–290.

Bergstrom, C. T., West, J. D., & Wiseman, M. A. (2008). The Eigenfactor metrics. *The Journal of Neuroscience* <http://doi.org/10.1523/JNEUROSCI.0003-08.2008> *Official Journal of the Society for Neuroscience*

Bernal, I. (2013). Open Access and the Changing Landscape of Research Impact Indicators: New Roles for Repositories. *Publications*, 1(2), 56–77. <http://doi.org/10.3390/publications1020056>

Bernal, I., & Pemau-Alonso, J. (2010). Statistics for repositories: a metric system for data on Digital.CSIC. *El Profesional de La Información*, 19(5), 534–543. <http://doi.org/10.3145/epi.2010.sep.15>

Bernius, S., & Hanauske, M. (2009). Open Access to Scientific Literature - Increasing Citations as an Incentive for Authors to Make Their Publications Freely Accessible. In *2009 42nd Hawaii International Conference on System Sciences* (pp. 1–9). <http://doi.org/10.1109/HICSS.2009.335>

- Bester, E. (2010). Les services pour les archives ouvertes: de la référence à l'expertise. *Documentaliste - Sciences de l'Information*, 47(4), 4–15.
- Bhardwaj, R. K. (2014). Institutional Repository Literature: A Bibliometric Analysis. *Science & Technology Libraries*, 33(2), 185–202.
- Bhat, M. H. (2010). Open Access Repositories: A Review. *Library Philosophy and Practice*.
- Biblioteca de la Universidad de Salamanca. (2011). 10 claves del acceso abierto. Retrieved from <https://s3.amazonaws.com/assets.rbl.ms/88643/origin.jpg>
- Björk, B.-C. (2007). A model of scientific communication as a global distributed information system. *Information Research – An International Electronic Journal*, 12(2).
- Björk, B.-C., Laakso, M., Welling, P., & Paetau, P. (2014). Anatomy of green open access. *Journal of the Association for Information Science and Technology*, 65(2). <http://doi.org/10.1002/asi.22963>
- Björk, B.-C., Welling, P., Laakso, M., Majlender, P., Hedlund, T., & Gudnason, G. (2010). Open access to the scientific journal literature: situation 2009. *PloS One*, 5(6), e11273. <http://doi.org/10.1371/journal.pone.0011273>
- BOAI. (2002). Budapest Open Access Initiative. Retrieved from <http://www.budapestopenaccessinitiative.org/>
- Bonilla-Calero, A. (2014). Institutional repositories as complementary tools to evaluate the quantity and quality of research outputs. *Library Review*, 63(1–2), 46–59. <http://doi.org/10.1108/LR-05-2013-0058>
- Boote, D., & Beile, P. (2005, September 1). Scholars before Researchers: On the Centrality of the Dissertation Literature Review in Research Preparation. *Educational Researcher*.
- Borgman, C. L., & Furner, J. (2005). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology*, 36(1), 2–72. <http://doi.org/10.1002/aris.1440360102>
- Borgoños Martínez, M. D. (2010). El movimiento Open Archives Initiative (OAI) y su repercusión en la difusión del Conocimiento. *Anales de Documentación*, 13, 23–40.
- Borrego, Á. (2015). Measuring compliance with a Spanish Government open access mandate. *Journal of the Association for Information Science and Technology*, 67(4), 757–764. <http://doi.org/10.1002/asi.23422>
- Breeding, M. (2015). *The Future of Library Resource Discovery: A white paper commissioned by the NISO Discovery to Delivery (D2D) Topic Committee*. Baltimore, MD.
- Brettle, A. (2003). Information skills training: a systematic review of the literature*. *Health Information and Libraries Journal*, 20(s1), 3–9. <http://doi.org/10.1046/j.1365-2532.20.s1.3.x>
- Brody, T. (2006). Evaluating research impact through open access to scholarly communication. Retrieved from <http://eprints.ecs.soton.ac.uk/13313>
- Brody, T., Harnad, S., & Carr, L. (2006). Earlier Web usage statistics as predictors of later citation

- impact. *Journal of the American Society for Information Science and Technology*, 57(8), 1060–1072. <http://doi.org/10.1002/asi.20373>
- Buehler, M. A., & Boateng, A. (2005). The evolving impact of institutional repositories on reference librarians. *Reference Services Review*, 33(3), 291–300. <http://doi.org/http://dx.doi.org/10.1108/00907320510611320>
- Burns, C. (2013). Google Scholar and free or open access scholarly content: impact on academic libraries. Retrieved from <https://mospace.umsystem.edu/xmlui/handle/10355/16178>
- Burns, C. S., Lana, A., & Budd, J. M. (2013). Institutional Repositories: Exploration of Costs and Value. *D-Lib Magazine*, 19(1/2). <http://doi.org/10.1045/january2013-burns>
- Calver, M. C., & Bradley, J. S. (2010). Patterns of citations of open access and non-open access conservation biology journal papers and book chapters. *Conservation Biology*, 24(3), 872–880. <http://doi.org/10.1111/j.1523-1739.2010.01509.x>
- Calzada Prado, F. J. (2010). *Repositorios, bibliotecas digitales y CRIAllos objetos de aprendizaje en la educación superior*. Buenos Aires, Argentina: Alfagrama Ediciones.
- Campbell, R., & Meadows, A. (2011). Scholarly journal publishing: where do we go from here? *Learned Publishing*, 24(3), 171–181. <http://doi.org/10.1087/20110305>
- Campbell, S. A., & Menk, D. W. (2003). Editor's introduction. *Review of Educational Research*, 73(2), 123–124. <http://doi.org/10.3102/00346543073002123>
- Carr, L., & Brody, T. (2007). Size Isn't Everything: Sustainable Repositories as Evidenced by Sustainable Deposit Profiles. *D-Lib Magazine*, 13(7–8).
- Carroll, M. W. (2013). Creative Commons and the openness of open access. *The New England Journal of Medicine*, 368(9), 789–91. <http://doi.org/10.1056/NEJMp1300040>
- Casany, M. J., Alier Forment, M., Conde, M. Á., & García-Peñalvo, F. J. (2009). SOA initiatives for eLearning. A Moodle case. In *23rd International Conference on Advanced Information Networking and Applications, AINA 2009, Workshops Proceedings. The International Symposium on Mining and Web (MAW 2009)* (pp. 750–755). Los Alamitos, California, USA: IEEE Computer Society.
- Casaus, A., & Ferreras-Fernández, T. (2009). GREDOS: Gestión del Repositorio Documental de la Universidad de Salamanca. In *IX Workshop REBIUN Proyectos digitales*. Salamanca, España. Retrieved from <http://sabus.usal.es/workshop/pdfs/4/gredos.pdf>
- Cassella, M. (2010). Institutional Repositories: an Internal and External Perspective on the Value of IRs for Researchers' Communities. *LIBER Quarterly*, 20(2), 210. <http://doi.org/10.18352/lq.7989>
- Center for Research Libraries. (2007). *TRAC Checklist is Published*. OCLC. Retrieved from http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf
- Cesare, R. Di, Luzi, D., & Ruggieri, R. (2008). The impact of Grey Literature in the web environment: A citation analysis using Google Scholar. *The Grey Journal*, 4(2), 83–96.

- Cesare, R. Di, & Ruggieri, R. (2006). The evaluation of Grey Literature using bibliometric indicators - A methodological proposal. In C. Anderson, KL and Thiery (Ed.), *Information for Responsible Fisheries: Libraries as Mediators, Proceedings* (pp. 57–63). Int Assoc Marine Science Libraries & Information Center.
- Chan, L. (2004). Supporting and Enhancing Scholarship in the Digital Age: The Role of Open Access Institutional Repository. *Canadian Journal of Communication*, 29(3).
- Chang, E., & West, M. (2006). Digital Ecosystems A Next Generation of the Collaborative Environment. In & I. K. I. G. Kotsis, D. Taniar, E. Pardede (Ed.), *Proceedings of iiWAS'2006 - The Eighth International Conference on Information Integration and Web-based Applications Services* (pp. 3–24). Yogyakarta, Indonesia.
- Chen, X. (2014). (2014). Open access in 2013: reaching the 50% milestone. *Serials Review*, 40(1). <http://doi.org/10.1080/00987913.2014.895556>
- Cheng, R., Cheung, S., & Kot, C. (2014). Explore ETD's' Visibility and Discoverability through Usage Analysis. In University of Leicester (Ed.), *ETD 2014*. Leicester, UK.
- Cheng, W. H., Ren, S. L., & Rousseau, R. (2012). Digital publishing and China's core scientific journals: a position paper. *Scientometrics*, 98(1), 11–22. <http://doi.org/10.1007/s11192-012-0873-8>
- Cheng, W., & Ren, S. (2008). Evolution of open access publishing in Chinese scientific journals. *Learned Publishing*, 21(2), 140+. <http://doi.org/10.1087/095315108X288884>
- Chillag, J. (1973). Don't be afraid of reports. *BLL Review*, 1(2), 39–51. <http://doi.org/10.1108/eb008432>
- Clobridge, A. (2014). All About Open Access Repositories: The Other Side of Open Access. *Online Searcher*, 38(5), 38–41.
- COAR. (2012). *The Current State of Open Access Repository Interoperability (2012)*. Retrieved from <https://www.coar-repositories.org>
- COAR. (2015). *COAR Roadmap Future Directions for Repository Interoperability*. Retrieved from <http://coar-repositories.org>
- COAR. Grupo de trabajo 2: Interoperabilidad del Repositorio. (2011). *El caso de Interoperabilidad para Repositorios de Acceso Abierto*. Retrieved from <https://www.coar-repositories.org/activities/repository-interoperability/coar-interoperability-roadmap/a-case-for-interoperability/>
- Cocciolo, A. (2010). Can Web 2.0 Enhance Community Participation in an Institutional Repository? The Case of PocketKnowledge at Teachers College, Columbia University. *The Journal of Academic Librarianship*, 36(4), 304–312. <http://doi.org/10.1016/j.acalib.2010.05.004>
- Codina, L. (19 octubre 2015). No lo llame Análisis Bibliográfico, llámelo Revisión Sistematizada. Y cómo llevarla a cabo con garantías: Systematized Reviews + SALSA Framework [Mensaje de blog]. Recuperado de <https://www.luiscodina.com/revison-sistematica-salsa-framework/>

- Conde, M. Á., García-Peñalvo, F. J., Casany, M. J., & Alier Forment, M. (2009). Adapting LMS architecture to the SOA: an Architectural Approach. In IEEE Computer Society (Ed.), *Proceedings of the Fourth International Conference on Internet and Web Applications and Services – ICIW 2009 (Venice/Mestre, Italy, 24-28 May 2009)* (pp. 322–327). Los Alamitos, California, USA.
- Conde, M. Á., García-Peñalvo, F. J., Rodríguez-Conde, M. J., Alier Forment, M., Casany, M. J., & Piguillem, J. (2014). An evolving Learning Management System for new educational environments using 2.0 tools. *Interactive Learning Environments*, 22(2), 188–204. <http://doi.org/10.1080/10494820.2012.745433>
- Conde González, M. Á., García-Peñalvo, F. J., Rodríguez-Conde, M. J., Alier Forment, M., & García-Holgado, A. (2014). Perceived openness of Learning Management Systems by students and teachers in education and technology courses. *Computers in Human Behavior*, 31, 517–526. <http://doi.org/10.1016/j.chb.2013.05.023>
- Connell, T., & Cetwinski, T. (2010). The Impact of Institutional Repositories on Technical Services. *Technical Services Quarterly*. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/07317131003765993>
- Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, 1(1), 104–126. <http://doi.org/10.1007/BF03177550>
- Cordes, R. (2004). Utilise-t-on la littérature grise? L'utilisation de l'analyse des citations pour mesurer l'influence du GESAMP, un organisme consul-tatif international des sciences marines. Is grey literature ever used? Using citation analysis to measure the impact of GE. *Canadian Journal of Information and Library Science*, 28(1), 49–69.
- Correia, A. M. R., & Neto, M. de C. (2002). The role of eprint archives in the access to, and dissemination of, scientific grey literature: LIZA - a case study by the National Library of Portugal. *Journal of Information Science*, 28(3), 231–241. <http://doi.org/10.1177/016555150202800305>
- Crow, R. (2002). *The Case for Institutional Repositories: A SPARC Position Paper*.
- Cullen, R., & Chawner, B. (2010). Institutional repositories: assessing their value to the academic community. *Performance Measurement and Metrics*, 11(2), 131–147. <http://doi.org/10.1108/14678041011064052>
- Dahlgren, P. (2005). The internet, public spheres, and political communication: dispersion and deliberation. *Political Communication*, (22), 147–162.
- Davis, P. M. (2009). Author-choice open-access publishing in the biological and medical literature: A citation analysis. *Journal of the American Society for Information Science and Technology*, 60(1), 3–8. <http://doi.org/10.1002/asi.20965>
- Davis, P. M. (2010). Does open access lead to increased readership and citations? A randomized controlled trial of articles published in aps journals. *Physiologist*, 53(6), 197–200.
- Davis, P. M. (2011). Open access, readership, citations: a randomized controlled trial of scientific journal publishing. *The FASEB Journal*, 25(7), 2129–2134. <http://doi.org/10.1096/fj.11-183988>

- Davis, P. M., & Connolly, M. J. L. (2007). Institutional Repositories: Evaluating the Reasons for Non-use of Cornell University's Installation of DSpace. *D-Lib Magazine*, 13(34).
- Davis, P. M., Lewenstein, B. V., Simon, D. H., Booth, J. G., & Connolly, M. J. L. (2008). Open access publishing, article downloads, and citations: randomised controlled trial. *British Medical Journal*, 337(7665), a568–a568. <http://doi.org/10.1136/bmj.a568>
- Delgado López-Cózar, E. (3 de enero de 2014). Google Scholar: presente y futuro, fortalezas y debilidades, reportaje de John Bohannon en Science [Mensaje de Blog]. Retrieved from <http://ec3noticias.blogspot.com.es/2014/01/google-scholar-wins-ravesbut-can-it-be.html>
- Delgado López-Cózar, E., Torres-Salinas, D., Jiménez-Contreras, E., & Ruiz-Pérez, R. (2006). Análisis bibliométrico y de redes sociales aplicado a las tesis bibliométricas defendidas en España (1976-2002): temas, escuelas científicas y redes académicas. *Revista Española de Documentación Científica*, 29(4), 493–524. <http://doi.org/10.3989/redc.2006.v29.i4.306>
- Deng, S., & Reese, T. (2009). Customized mapping and metadata transfer from DSpace to OCLC to improve ETD work flow. *New Library World*, 110(5/6), 249–254. <http://dx.doi.org/10.1108/03074800910954271>
- Deng, Z., & Li, Q. (2008). Upgrading the User Experience of Digital Library. In *2008 International Conference on Computer Science and Software Engineering* (Vol. 5, pp. 832–835). IEEE. <http://doi.org/10.1109/CSSE.2008.184>
- Donovan, J. M., & Watson, C. A. (2011). Citation advantage of open access legal scholarship. *Law Library Journal*, 103(4), 553–573.
- Donovan, J. M., Watson, C. A., & Osborne, C. (2014). The Open Access Advantage for American Law Reviews. *SSRN Electronic Journal*. <http://doi.org/10.2139/ssrn.2506913>
- Dorner, D. G., & Revell, J. (2012). Subject librarians' perceptions of institutional repositories as an information resource. *Online Information Review*, 36(2), 261–277. <http://doi.org/10.1108/14684521211229066>
- Dowling, T., & Steans, R. (2011). Managing Consortial ETD Repositories. In *USETDA 2011: The Magic of ETDs... Where Creative Minds Meet, May 18-20*. Orlando, Florida.
- Ducloy, J., Ducasse, J. P., Foulonneau, M., Grivel, L., Hénaff, D. Le, & Nicolas, Y. (2006). Metadata towards an e-research cyberinfrastructure: the case of French PhD theses. In *DCMI '06: Proceedings of the 2006 international conference on Dublin Core and Metadata Applications* (pp. 133–148).
- Dunn, K. (2005). Impact of the inclusion of grey literature on the scholarly communication patterns of an interdisciplinary specialty. In Farace, DJ (Ed.), *GL6: Work on Grey in Progress, Conference Proceedings* (pp. 50–54).
- Echeverria, M., Stuart, D., & Blanke, T. (2015). Medical theses and derivative articles: Dissemination of contents and publication patterns. *Scientometrics*, 102(1), 559–586. <http://doi.org/10.1007/s11192-014-1442-0>
- Egghe, L. (2005). Expansion of the field of informetrics: Origins and consequences. *Information Processing & Management*. <http://doi.org/10.1016/j.ipm.2005.03.011>

- Elsevier. (2013). *Performance of the UK research base: international comparison - 2013. BIS research papers*. Retrieved from <https://www.gov.uk/government/publications/performance-of-the-uk-research-base-international-comparison-2013>
- Erren, T. C., Cullen, P., & Erren, M. (2009). How to surf today's information tsunami: on the craft of effective reading. *Medical Hypotheses*, 73(3), 278–9. <http://doi.org/10.1016/j.mehy.2009.05.002>
- Evans, J. A., & Reimer, J. (2009). Open access and global participation in science. *Science (New York, N.Y.)*, 323(5917), 1025.
- Executive Office of the President National Science and Technology Council. (2012). *Interagency Public Access Coordination: A report to Congress on the Coordination of Policies related to dissemination and long-term stewardship of Federally Scientific Research*. Retrieved from https://www.whitehouse.gov/sites/default/files/microsites/ostp/public_access-final.pdf
- Eysenbach, G. (2006). Citation Advantage of Open Access Articles. *PLoS Biology*, 4(5), e157. <http://doi.org/10.1371/journal.pbio.0040157>
- Ezema, I. (2011). Building open access institutional repositories for global visibility of Nigerian scholarly publication. *Library Review* 60(6), 473-485.
- Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2007). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *The FASEB Journal*, 22(2), 338–342. <http://doi.org/10.1096/fj.07-9492LSF>
- Fan, W. (2015). Contribution of the institutional repositories of the Chinese Academy of Sciences to the webometric indicators of their home institutions. *Scientometrics*, 105(3), 1889–1909. article. <http://doi.org/10.1007/s11192-015-1758-4>
- Farace, D. J., Frantzen, J., Schöpfel, J., Stock, C., & Boekhorst, A. K. (2006). Access to Grey Content: An Analysis of Grey Literature based on Citation and Survey Data: A Follow-up Study. In J. Farace, DJ and Frantzen (Ed.), *Seventh International Conference On Grey Literature, G17 Conference Proceedings* (pp. 194–203).
- FECYT. (2014). *Recomendaciones para la implementación del artículo 37 Difusión en Acceso Abierto de la Ley de la Ciencia, la Tecnología y la Innovación*.
- Fenner, M., & Lin, J. (2014). Novel Research Impact Indicators. *Liber Quarterly: The Journal of European Research Libraries*, 23(4), 300–309.
- Ferreira, M., Rodrigues, E., Baptista, A. A., & Saraiva, R. (2008). Carrots and Sticks. *D-Lib Magazine*, 14(1/2). <http://doi.org/10.1045/january2008-ferreira>
- Ferreras-Fernández, T. (2008). *Dublin Core Cualificado: documento de trabajo*. Retrieved from <http://hdl.handle.net/10366/56268>
- Ferreras-Fernández, T. (2010a). Las bibliotecas universitarias y el acceso abierto a la comunicación científica. In *4as. Jornadas BUCLE sobre Bibliotecas Universitarias*. Ávila. Retrieved from <http://hdl.handle.net/10366/76793>

- Ferreras-Fernández, T. (2010b). *Manual de usuario Gredos: edición de registros; plantillas, creación y tipología*. Retrieved from <http://hdl.handle.net/10366/82870>
- Ferreras-Fernández, T. (2010c). *Preservación digital en repositorios institucionales: Gredos*. Retrieved from <http://hdl.handle.net/10366/83130>
- Ferreras-Fernández, T. (2011a). Acceso abierto al conocimiento, políticas gubernamentales y derechos de autor: editorial. *Boletín SCOPEO*, (41).
- Ferreras-Fernández, T. (2011b). El alcance internacional del Acceso Abierto. Retrieved from http://eprints.rclis.org/handle/10760/16352#.T3l6xH6_t1s.mendeley
- Ferreras-Fernández, T. (2013a). El Repositorio Gredos de la Universidad de Salamanca como proveedor de datos de Europea. *Boletín de La ANABAD*, 63(3), 407–420.
- Ferreras-Fernández, T. (2013b). La interoperabilidad: el pegamento técnico para conectar repositorios. *Blok de bid*.
- Ferreras-Fernández, T. (2015). *Plan de investigación: Visibilidad e impacto de la literatura gris científica en repositorios institucionales de acceso abierto. Estudio de caso bibliométrico del repositorio Gredos de la Universidad de Salamanca*. Retrieved from <http://hdl.handle.net/10366/125851>
- Ferreras-Fernández, T., & Casaus-Peña, A. (2008). *Memoria de Trabajo del Repositorio Institucional de la Universidad de Salamanca*. Retrieved from <http://hdl.handle.net/10366/123014>
- Ferreras-Fernández, T., García-Peñalvo, F. J., & Merlo-Vega, J. A. (2015). Open access repositories as channel of publication scientific grey literature. In *Proceedings of the 3rd International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM '15* (pp. 419–426). New York, New York, USA: ACM Press. <http://doi.org/10.1145/2808580.2808643>
- Ferreras-Fernández, T., & Merlo-Vega, J. A. (2010). Políticas de promoción del acceso abierto en la Universidad de Salamanca. In *4as. Jornadas de OS-Repositorios*.
- Ferreras-Fernández, T., & Merlo-Vega, J. A. (2015). Repositorios de acceso abierto: un nuevo modelo de comunicación científica. *La Revista de la Sociedad ORL CLCR en el repositorio Gredos. Rev Soc Otorrinolaringol Castilla Leon Cantab La Rioja*, 6(12), 94–113.
- Ferreras-Fernández, T., Merlo-Vega, J. A., & García-Peñalvo, F. J. (2013a). Impact of scientific content in open access institutional repositories. In Francisco José García-Peñalvo (Ed.), *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM '13) (Salamanca, Spain, November 14-15, 2013)* (pp. 357–363). New York, USA: ACM Press. <http://doi.org/10.1145/2536536.2536590>
- Ferreras-Fernández, T., Merlo-Vega, J. A., & García-Peñalvo, F. J. (2013b). Science 2.0 supported by open access repositories and open linked data. In *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM '13) (Salamanca, Spain, November 14-15, 2013)* (pp. 331–332). New York, New York, USA: ACM Press. <http://doi.org/10.1145/2536536.2536586>

- Ferreras-Fernández, T., Merlo-Vega, J. A., & García-Peñalvo, F. J. (2015). Visibilidad de la literatura gris científica a través de repositorios. El caso de las tesis doctorales en Gredos. In *XIV Workshop de REBIUN sobre proyectos digitales y VI Jornadas de OS-Repositorios (11-13 de marzo de 2015, Córdoba, Spain)*. Retrieved from <http://hdl.handle.net/10396/12630>
- Fichman, R. G. (1992). Information technology diffusion: A review of empirical research. In *Proceedings of the 13th International Conference on Information Systems* (pp. 195–206).
- Fidalgo-Blanco, Á., Sein-Echaluce Lacleta, M. L., Borrás Gené, O., & García-Peñalvo, F. J. (2014). Educación en abierto: Integración de un MOOC con una asignatura académica. *Education in the Knowledge Society (EKS)*, 15(3), 233–255.
- Finch, D. J., Bell, S., Bellingan, L., Campbell, R., Donnelly, P., Gardner, R., ... Jubb, M. (2013). Accessibility, sustainability, excellence: How to expand access to research publications. Executive summary. *International Microbiology*, 16(2). <http://doi.org/10.2436/20.1501.01.187>
- Foster, N. F., & Gibbons, S. (2005). Understanding Faculty to Improve Content Recruitment for Institutional Repositories. *D-Lib Magazine*, 11(1).
- Fox, E. A. (2014). Improving the ETD landscape. In *ETD2014: NDLTD's 17th International Symposium on Electronic Theses & Dissertations*. Leicester, UK: University of Leicester.
- Frandsen, T. F. (2009). The integration of open access journals in the scholarly communication system: Three science fields. *Information Processing & Management*, 45(1), 131–141. <http://doi.org/10.1016/j.ipm.2008.06.001>
- Friend, F. (2011). *Open Access Business Models for Research Funders and Universities*. Copenhagen. Retrieved from <http://www.oerafrica.org/resource/open-access-business-models-research-funders-and-universities>
- Frisch, N. K., Nathan, R., Ahmed, Y. K., Shidham, V. B., & Ahmed, Y. K. (2014). Authors attain comparable or slightly higher rates of citation publishing in an open access journal (CytoJournal) compared to traditional cytopathology journals - A five year (2007-2011) experience. *CytoJournal*, 11(1), 10. <http://doi.org/10.4103/1742-6413.131739>
- Galina Russell, I. (2009). *Electronic resources and institutional repositories in informal scholarly communication and publishing. Doctoral thesis, UCL (University College London)*. UCL (University College London).
- Galina Russell, I. (2011). La visibilidad de los recursos académicos: Una revisión crítica del papel de los repositorios institucionales y el acceso abierto. *Investigación bibliotecológica*, 25(53), 159–183.
- Galindo, F. M., & Llopes, I. R. (2015). Publicar en acceso abierto con impacto. Retrieved from <https://riunet.upv.es/handle/10251/50453>
- Galloway, L. M., & Pease, J. L. (2013). Altmetrics for the Information Professional: A Primer. *Libraries' and Librarians' Publications, Paper 105*(June), 1–5.
- García Peñalvo, F. J., Rodríguez Conde, M. J., Seoane Pardo, A. M., Conde González, M. Á.,

- Zangrando, V., & García Holgado, A. (2012). GRIAL (GRupo de investigación en InterAcción y eLearning), USAL. *IE Comunicaciones: Revista Iberoamericana de Informática Educativa*, 15, 85–94.
- García-Holgado, A., & García-Peñalvo, F. J. (2013a). Análisis de integración de soluciones basadas en software como servicio para la implantación de ecosistemas tecnológicos corporativos. In R. V. R. J. Cruz Benito, A. García Holgado, S. García Sánchez, D. Hernández Alfageme, M^a Navarro Cáceres (Ed.), *Avances en Informática y Automática. 7^o Workshop* (pp. 55–72). Salamanca: Departamento de Informática y Automática de la Universidad de Salamanca.
- García-Holgado, A., & García-Peñalvo, F. J. (2013b). The evolution of the technological ecosystems. In *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality - TEEM '13* (pp. 565–571). New York, New York, USA: ACM Press. <http://doi.org/10.1145/2536536.2536623>
- García-Holgado, A., & García-Peñalvo, F. J. (2016). Architectural pattern to improve the definition and implementation of eLearning ecosystems. *Science of Computer Programming*, 129, 20–34. <http://doi.org/10.1016/j.scico.2016.03.010>
- García-Peñalvo, F. J. (2008a). Docencia. In Laviña Orueta & L. Mengual Pavón (Ed.), *Libro Blanco de la Universidad Digital 2010* (pp. 29–61). Barcelona, España: Ariel.
- García-Peñalvo, F. J. (2008b). El conocimiento abierto en español en la Universidad. Ponencia presentada Seminario “Pensar en español” celebrado en Cartagena de Indias (Colombia). Retrieved from <http://hdl.handle.net/10366/56045>
- García-Peñalvo, F. J. (2008c). Un nuevo paradigma de universidad: La universidad digital. In *III Congreso Internacional, Software libre y Web 2.0 Educación y Formación Universidad del País Vasco, Campus de Leioa (Vizcaya)*. Retrieved from <http://repositorio.grial.eu/handle/grial/572>
- García-Peñalvo, F. J. (2008d). *Advances in E-Learning: Experiences and Methodologies*. Hershey, PA, USA: Information Science Reference (formerly Idea Group Reference).
- García-Peñalvo, F. J. (2009). La Tecnología en la Universidad de Salamanca. In *Jóvenes con la Cultura, la Ciencia y la Tecnología (II ed.)*. Campamento de verano: Tecnología y Universidad: la ciencia en nuestras manos, Salamanca. Salamanca: Fundación Universidades de Castilla y León (JCyL) / Universidad de Salamanca.
- García-Peñalvo, F. J. (2011). La Universidad de la próxima década: la Universidad Digital. In C. Suárez-Guerrero & F. J. García-Peñalvo (Ed.), *Universidad y Desarrollo Social de la Web* (pp. 181–187). Washington, DC: Editandum.
- García-Peñalvo, F. J. (2013a). Aportaciones de la Ingeniería en una Perspectiva Multicultural de la Sociedad del Conocimiento. *VAEP-RITA*, 1(4), 201–202.
- García-Peñalvo, F. J. (2013b). Education in knowledge society: A new PhD programme approach. In F. J. García-Peñalvo (Ed.), *Proceedings of the First International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'13) (Salamanca, Spain, November 14-15, 2013)* (pp. 575–577). New York, NY, USA: ACM.

- García-Peñalvo, F. J. (2013c). Formación en la sociedad del conocimiento, un programa de doctorado con una perspectiva interdisciplinar. *Revista Teoría de La Educación: Educación Y Cultura En La Sociedad de La Información*, 15(1), 4–9.
- García-Peñalvo, F. J. (2015). Ecosistemas tecnológicos. *IEEE VAEP-RITA*, 3(1), 36–37. Retrieved from <http://hdl.handle.net/10366/125948>
- García-Peñalvo, F. J. (2015). Engineering contributions to a Knowledge Society multicultural perspective. *IEEE Revista Iberoamericana de Tecnologías Del Aprendizaje (IEEE RITA)*, 10(1), 17–18. <http://doi.org/10.1109/RITA.2015.2391371>
- García-Peñalvo, F. J. (2016a). Digital Humanities Data Processing. *Journal of Information Technology Research*, 9.
- García-Peñalvo, F. J. (2016b). Diseminación y divulgación científica. Retrieved from <http://gredos.usal.es/jspui/handle/10366/127252>
- García-Peñalvo, F. J., Alier Forment, M., & Lytras, M. D. (2012). Some Reflections about Service Oriented Architectures, Cloud Computing Applications, Services and Interoperability. *Journal of Universal Computer Science*, 18(11), 1405–1409.
- García-Peñalvo, F. J., Conde, M. Á., Alier Forment, M., & Casany, M. J. (2011). Opening Learning Management Systems to Personal Learning Environments. *Journal of Universal Computer Science*, 17(9), 1222–1240. <http://doi.org/doi:10.3217/jucs-017-09-1222>
- García-Peñalvo, F. J., García de Figuerola, C., & Merlo-Vega, J. A. (2010a). Open knowledge: challenges and facts. *Online Information Review*, 34(4), 520–539. <http://doi.org/10.1108/14684521011072963>
- García-Peñalvo, F. J., García de Figuerola, C., & Merlo-Vega, J. A. (2010b). Open knowledge management in higher education. *Online Information Review*, 34(4), 517–519.
- García-Peñalvo, F. J., Merlo-Vega, J. A., Ferreras-Fernández, T., Casaus-Peña, A., Albás-Aso, L., & Atienza-Díaz, M. L. (2010). Qualified Dublin Core Metadata Best Practices for GREDOS. *Journal of Library Metadata*, 10(1), 13–36. <http://doi.org/10.1080/19386380903546976>
- García-Peñalvo, F. J., & Rivera, S. M. (2009). Digital University 2010. In & P. J. R. J. R. Canay, J. Franco (Ed.), *Key of the European Space of Knowledge. (Santiago de Compostela, Spain, June 23rd - 26th, 2009)* (p. 32). Santiago de Compostela, Spain: Universidade de Santiago de Compostela.
- García-Peñalvo, F. J., & Seoane-Pardo, A. M. (2015). Una revisión actualizada del concepto de eLearning. Décimo Aniversario. *Education in the Knowledge Society*, 16(1), 119–144. <http://doi.org/http://dx.doi.org/10.14201/eks2015161119144>
- Garfield, E. (1955). Citation indexes to science: a new dimension in documentation through association of ideas. *Science*, 122, 108–111.
- Garfield, E. (1970). Citation indexing for studying science. *Nature*, (227), 669–771.
- Garfield, E. (2006). The History and Meaning of the Journal Impact Factor. *JAMA*, 295(1), 90.

<http://doi.org/10.1001/jama.295.1.90>

- Gargouri, Y., Hajjem, C., Larivière, V., Gingras, Y., Carr, L., Brody, T., & Harnad, S. (2010). Self-selected or mandated, open access increases citation impact for higher quality research. *PLoS ONE*, *5*(10), e13636. <http://doi.org/10.1371/journal.pone.0013636>
- Gargouri, Y., Lariviere, V., Gingras, Y., Brody, T., Carr, L., & Harnad, S. (2012). Testing the Finch Hypothesis on Green OA Mandate Ineffectiveness, 6. Digital Libraries. Retrieved from <http://arxiv.org/abs/1210.8174>
- Gargouri, Y., Larivière, V., Gingras, Y., Carr, L., & Harnad, S. (2012). Green and Gold Open Access Percentages and Growth, by Discipline. Digital Libraries. Retrieved from <http://arxiv.org/abs/1206.3664>
- Gaulé, P., & Maystre, N. (2011). Getting cited: Does open access help? *Research Policy*, *40*(10), 1332–1338. <http://doi.org/10.1016/j.respol.2011.05.025>
- Gentil-Beccot, A., Mele, S., & Brooks, T. C. (2010). Citing and reading behaviours in high-energy physics. *Scientometrics*, *84*(2), 345–355. <http://doi.org/10.1007/s11192-009-0111-1>
- Gherab Martin, K. J. (2015). El Informe Finch y el acceso abierto a las publicaciones biomédicas. *El Profesional de La Información*, *24*(5), 631–639.
- Ghosh, M. (2009). E-theses and Indian academia: A case study of nine ETD digital libraries and formulation of policies for a national service. *The International Information & Library Review*, *41*(1), 21–33. <http://doi.org/10.1016/j.iilr.2008.08.002>
- Giesecke, J. (2011). Institutional Repositories: Keys to Success. *Journal of Library Administration*, *51*(5–6), 529–542. <http://doi.org/10.1080/01930826.2011.589340>
- Giglia, E. (2010). The impact factor of open access journals: data and trends. Retrieved from <http://eprints.rclis.org/14666>
- Giusti, M. De. (2014). Visibilidad e impacto de la producción de una institución a través de los repositorios de acceso abierto. *XIII Congreso Nacional de Bibliotecología Y XX ...* Retrieved from <http://sedici.unlp.edu.ar/handle/10915/41101>
- González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). A new approach to the metric of journals' scientific prestige: The SJR indicator. *Journal of Informetrics*, *4*(3), 379–391. <http://doi.org/10.1016/j.joi.2010.03.002>
- Graham, T. W. (2000). Scholarly Communication. *Serials: The Journal for the Serials Community*, *13*(1), 3–11. <http://doi.org/10.1629/133>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, *26*(2), 91–108. <http://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Guédon, J.-C. (2008a). Mixing and Matching the Green and Gold Roads to Open Access—Take 2. *Serials Review*, *34*(1), 41–51. <http://doi.org/10.1016/j.serrev.2007.12.008>
- Guédon, J.-C. (2008b, March 3). Accès libre, archives ouvertes et États-nations : les stratégies

- du possible. Retrieved from http://archivesic.ccsd.cnrs.fr/sic_00277755/en/
- Gulley, N. (2013). Creative Commons: challenges and solutions for researchers; a publisher's perspective of copyright in an open access environment. *Insights: The UKSG Journal*, 26(2), 168–173.
- Hagen, J. H. (2007). Building Effective Discovery Tools for Academic Promotion and Tenure Evidence: the Added Value of ETD and Institutional Repository Metadata, Citation and Access. In *ETD 2007 10th International*. Uppsala, Sweden.
- Hajjem, C., Harnad, S., & Gingras, Y. (2005). Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact. *IEEE Data Engineering Bulletin*, 28(4), 39–47. Retrieved from <http://arxiv.org/abs/cs/0606079>
- Hakimjavadi, H., & Masrek, M. N. (2013). Evaluation of interoperability protocols in repositories of electronic theses and dissertations. *Program: Electronic Library and Information Systems*, 47(1), 34–59. <http://doi.org/10.1108/00330331211296303>
- Halbert, M. (2008). Integrating ETD Services into Campus Institutional Repository Infrastructures Using Fedora. In *Proceedings of Tenth International Symposium on Electronic Theses and Dissertations*. Uppsala University Library NDLTD.
- Halliday, L. (2001). Scholarly communication, scholarly publication and the status of emerging formats. *Information Research*, 6(4).
- Hampton, S. E., & Parker, J. N. (2011). Collaboration and Productivity in Scientific Synthesis. *BioScience*, 61(11), 900–910. <http://doi.org/10.1525/bio.2011.61.11.9>
- Hanief Bhat, M. (2010). Interoperability of open access repositories in computer science and IT – an evaluation. *Library Hi Tech*, 28(1), 107–118. <http://doi.org/10.1108/07378831011026724>
- Harnad, S. (2001). The self-archiving initiative. *Nature*, (410), 1024–1025.
- Harnad, S. (2006). Maximizing research impact through institutional and national open-access self-archiving mandates. In *CRIS 2006: Enabling Interaction and Quality: Beyond the Hanseatic League - 8th International Conference on Current Research Information Systems*. Canada Research Department in Cognitive Sciences, Université du Québec à Montréal, Canada: Leuven University Press.
- Harnad, S. (2008). The University's Mandate To Mandate Open Access. Retrieved from <http://eprints.ecs.soton.ac.uk/16831>
- Harnad, S. (2009). Self-Archiving, Metrics, and Mandates. *European Science Editing*, 35(1), 6–8.
- Harnad, S. (2011). Open Access to Research: Changing Researcher Behavior Through University and Funder Mandates. *JEDEM Journal of Democracy and Open Government*, 3(1), 33–41.
- Harnad, S. (2012). Against raising green OA goalpost from gratis to CC-BY. Retrieved May 7, 2016, from <http://openaccess.eprints.org/index.php?/archives/946-Against-Raising-Green-OA-Goalpost-From-Gratis-To-CC-BY.html>

- Harnad, S. (2012a). Finch Fiasco in Figures. Retrieved March 25, 2015, from <https://groups.google.com/a/arl.org/forum/#!topic/sparc-oaforum/KainrZup9Os>
- Harnad, S. (2012b, July 4). Why the UK Should Not Heed the Finch Report. *LSE Impact of Social Sciences Blog*. Retrieved from <http://eprints.soton.ac.uk/341128/1/LSEFinch.pdf>
- Harnad, S. (2013a). Finch II: "Our Mind's Made Up: Don't Confuse Us With Facts." Retrieved March 30, 2016, from <http://openaccess.eprints.org/index.php?/archives/1074-Finch-II-Our-Minds-Made-Up-Dont-Confuse-Us-With-Facts.html>
- Harnad, S. (2013b). Harnad Comments on HEFCE/REF Open Access Mandate Proposal. *Open Access and Submissions to the REF Post-2014*. Retrieved from <http://eprints.soton.ac.uk/349893>
- Harnad, S., & Brody, T. (2004). Comparing the impact of Open Access (OA) vs. non-OA articles in the same journals. *D-Lib Magazine*, 10(6). <http://doi.org/10.1045/june2004-harnad>
- Harnad, S., Brody, T., Vallières, F., Carr, L., Hitchcock, S., Gingras, Y., ... Hilf, E. R. (2008). The Access/Impact Problem and the Green and Gold Roads to Open Access: An Update. *Serials Review*, 34(1), 36–40. <http://doi.org/10.1016/j.serrev.2007.12.005>
- Harnad, S., Carr, L., & Swan, A. (2009). Open Access Repositories-maximizing and measuring research impact through university and research-funder open-access self-archiving mandates. Retrieved from <http://eprints.utas.edu.au/9273/>
- Harzing, A.-W., & Van der Wal, R. (2007). Google Scholar: the democratization of citation analysis. *Ethics in Science and Environmental Politics*, 8(1).
- Haustein, S., & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. *Journal of Informetrics*, 5(3), 446–457. <http://doi.org/10.1016/j.joi.2011.04.002>
- Hawkins, A. R., Kimball, M. A., & Ives, M. (2013). Mandatory Open Access Publishing for Electronic Theses and Dissertations: Ethics and Enthusiasm. *The Journal of Academic Librarianship*, 39(1), 32–60. <http://doi.org/10.1016/j.acalib.2012.12.003>
- Heyse, W. (2007). Supporting change and diversity for theses in an institutional repository. In *ETD 2007 10th International Symposium on Electronic Theses and Dissertations*. Uppsala, Sweden. Retrieved from <http://epc.uu.se/etd2007/files/papers/paper-46.pdf>
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569–72. <http://doi.org/10.1073/pnas.0507655102>
- Ho, A. (2016). Scholarly Communication and Research Libraries. *Library Presentations*. Retrieved from http://uknowledge.uky.edu/libraries_present/101
- Hood, W. W., & Wilson, C. S. (2001). The Literature of Bibliometrics, Scientometrics, and Informetrics. *Scientometrics*, 52(2), 291–314. <http://doi.org/10.1023/A:1017919924342>
- Houghton, J., & Sheehan, P. (2006, July 1). The Economic Impact of Enhanced Access to Research Findings. *CSES Working Paper*.

- Houghton, J., & Sheehan, P. (2009). Estimating the Potential Impacts of Open Access to Research Findings. *Economic Analysis and Policy*, 39(1), 127–142. [http://doi.org/10.1016/S0313-5926\(09\)50048-3](http://doi.org/10.1016/S0313-5926(09)50048-3)
- Houghton, J. W. (2009). Open Access: What are the Economic Benefits? A Comparison of the United Kingdom, Netherlands and Denmark. *SSRN Electronic Journal*. <http://doi.org/10.2139/ssrn.1492578>
- Houghton, J. W. (2010). Economic Implications of Alternative Publishing Models: Self-archiving and Repositories. *LIBER Quarterly*, 19(3–4), 275. <http://doi.org/10.18352/lq.7966>
- Houghton, J. W., & Oppenheim, C. (2010). The economic implications of alternative publishing models. *Prometheus*, 28(1), 41–54. <http://doi.org/10.1080/08109021003676359>
- Houghton, J. W., Rasmussen, B., Sheehan, P., Oppenheim, C., Morris, A., Creaser, C., ... Gourlay, A. (2009). *Economic implications of alternative scholarly publishing models: Exploring the costs and benefits*. Jisc.
- Houghton, J. W., & Swan, A. (2013). Planting the Green Seeds for a Golden Harvest: Comments and Clarifications on “Going for Gold.” *D-Lib Magazine*, 19(1/2).
- Hrynaszkiewicz, I., Busch, S., & Cockerill, M. J. (2013). Licensing the future: report on BioMed Central’s public consultation on open data in peer-reviewed journals. *BMC Research Notes*, 6, 318. <http://doi.org/10.1186/1756-0500-6-318>
- Huaroto, L. (2008). Cybertesis as a cooperative process for the implementation of the Digital Thesis Peruvian Network. In The Robert Gordon University (Ed.), *ETD 2008 11th International Symposium on Electronic Theses and Dissertations*. Aberdeen, UK.
- Illanas, A., & Llorens Largo, F. (2011). Los retos Web 2.0 de cara al EEES. In C. Suárez Guerrero y F. J. García Peñalvo (Ed.), *Universidad y Desarrollo Social de la Web* (pp. 13–34). Washington, D.C.: Editandum.
- Ingwersen, P., & Elleby, A. (2011). Do Open Access Working Papers Attract more Citations Compared to Printed Journal Articles from the same Research Unit? In Noyons, E and Ngulube, P and Leta, J (Ed.), *Proceedings Of Issi 2011: The 13th Conference of The International Society For Scientometrics And Informetrics, vols 1 and 2* (pp. 327–332).
- ISO. (1986). *International Standard 7144, presentation of theses and similar documents*. Geneva: International organization for standardization, documentation and Information.
- Jacsó, P. (2012). Google Scholar Metrics for Publications. *Online Information Review*, 36(4), 604–619. <http://doi.org/10.1108/14684521211254121>
- Jain, P. (2011). New trends and future applications/directions of institutional repositories in academic institutions. *Library Review*, 60(2), 125–141.
- Jo, K., & Tsuchide, I. (2013). Report of the COAR (Confederation of Open Access Repositories) Third Annual Meeting 2012. *Journal of College and University Libraries*, (97), 56–64.
- Johnson, R. (2015). *Making Open Access Work for Authors, Institutions and Publishers: A Report on an Open Access Roundtable Hosted by the Copyright Clearance Center Inc, Copyright*

Clearance Center, Danvers, MA.

Johnson, R. K. (2002). Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication. *D-Lib Magazine*, 8(11).

Joint Information Systems Committee. (2014). How Jisc is helping researchers.

Joint Information Systems Committee (JISC). (2008). *Managing and Sharing research Resources: How Open Access Repositories Can Help*.

Jones, R., & Andrew, T. (2005). Open access, open source and e-theses: the development of the Edinburgh Research Archive. *Program: Electronic Library and Information Systems*, 39(3), 198–212. <http://doi.org/10.1108/00330330510610555>

Joseph, D., Ng, K.-Y., Koh, C., & Ang, S. (2007). Turnover of information technology professionals: a narrative review, meta-analytic structural equation modeling, and model development. *MIS Quarterly*, 31(3), 547–577.

Juznic, P. (2010). Grey Literature produced and published by Universities: A Case for ETDs. In D. Farace & J. Schöpfel (Eds.), *Grey Literature in Library and Information Studies* (pp. 39–51). De Gruyter Saur.

Kanjilal, U., & Das, A. K. (2015). *Introduction to open access*. (UNESCO, Ed.). Paris.

Katz, J. E., Rice, R. E., & Aspden, P. (2001). The Internet, 1995-2000: Access, Civic Involvement, and Social Interaction. *American Behavioral Scientist*, 45(3), 405–419. <http://doi.org/10.1177/0002764201045003004>

Keefer, A. (2007, February 12). Los repositorios digitales universitarios y los autores. *Anales de Documentación*. <http://doi.org/10.6018/analesdoc.10.0.1151>

Kennan, M. A., & Wilson, C. (2006). Institutional repositories: review and an information systems perspective. *Library Management*, 27(4/5), 236–248. <http://doi.org/10.1108/01435120610668179>

Ketcham, C. M., & Crawford, J. M. (2007). The impact of review articles. *Laboratory Investigation; a Journal of Technical Methods and Pathology*, 87(12), 1174–85. <http://doi.org/10.1038/labinvest.3700688>

Kim, H. H., & Kim, Y. H. (2008). Usability study of digital institutional repositories. *The Electronic Library*, 26(6), 863–881. <http://doi.org/10.1108/02640470810921637>

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, 43(6), 740–755. <http://doi.org/10.1016/j.im.2006.05.003>

Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering (EBSE Technical Report EBSE-2007-01)*.

Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, 51(1), 7–15. <http://doi.org/10.1016/j.infsof.2008.09.009>

- Koler-Povh, T., Juznic, P., & Turk, G. (2014). Impact of open access on citation of scholarly publications in the field of civil engineering. *Scientometrics*, 98(2), 1033–1045. <http://doi.org/10.1007/s11192-013-1101-x>
- Koler-Povh, T., Mikos, M., & Turk, G. (2014). Institutional repository as an important part of scholarly communication. *Library Hi Tech*, 32(3), 423–434. <http://doi.org/10.1108/LHT-10-2013-0146>
- Konkie, S., & Scherer, D. (2013). New Opportunities for Repositories in the Age of Altmetrics. *Bulletin ASIST*, (April/May).
- Koufogiannakis, D., & Crumley, E. (2006). Research in librarianship: issues to consider. *Library Hi Tech*, 24(3), 324–340. <http://doi.org/10.1108/07378830610692109>
- Kousha, K., & Abdoli, M. (2010). The citation impact of Open Access agricultural research A comparison between OA and non-OA publications. *Online Information Review*, 34(5), 772–785. <http://doi.org/10.1108/14684521011084618>
- Kousha, K., & Thelwall, M. (2007a). How is science cited on the web? A classification of google unique web citations. *Journal of the American Society for Information Science and Technology*, 58(11), 1631–1644. <http://doi.org/10.1002/asi.20649>
- Kousha, K., & Thelwall, M. (2007b). Sources of Google Scholar citations outside the Science Citation Index: A comparison between four science disciplines. *Scientometrics*, 74(2), 273–294. <http://doi.org/10.1007/s11192-008-0217-x>
- Kroth, P. J., Phillips, H. E., & Hannigan, G. G. (2010). Institutional Repository Access Patterns of Nontraditionally Published Academic Content: What Types of Content Are Accessed the Most? *Journal of Electronic Resources in Medical Libraries*, 7(3), 189–195. <http://doi.org/10.1080/15424065.2010.505515>
- Kullman, L. (2014). The Effect of Open Access on Citation Rates os Self-archived Articles at Chalmers. In *IATUL Annual Conference Proceedings* (pp. 1–5).
- Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., Henneken, E., & Murray, S. S. (2005). The effect of use and access on citations. *Information Processing & Management*, 41(6), 1395–1402. <http://doi.org/10.1016/j.ipm.2005.03.010>
- Kyrillidou, M., & Young, M. (2008). ARL Statistics 2005-2006. Retrieved from <http://publications.arl.org/ARL-Statistics-2005-2006/13>
- Laakso, M. (2014). Green open access policies of scholarly journal publishers: a study of what, when, and where self-archiving is allowed. *Scientometrics*, 99(2), 475–494. <http://doi.org/10.1007/s11192-013-1205-3>
- Laakso, M., & Björk, B.-C. (2013). Delayed open access: An overlooked high-impact category of openly available scientific literature. *Journal of the American Society for Information Science and Technology*, 64(7), 1323–1329. <http://doi.org/10.1002/asi.22856>
- Lagoze, C., & Van de Sompel, H. (2003). The making of the Open Archives Initiative Protocol for Metadata Harvesting. *Library Hi Tech*, 21(2), 118–128. <http://doi.org/10.1108/07378830310479776>

- Lagzian, F., Abrizah, A., & Wee, M.-C. (2015). Measuring the gap between perceived importance and actual performance of institutional repositories. *Library & Information Science Research*, 37(2), 147–155. <http://doi.org/10.1016/j.lisr.2014.06.007>
- Lambert, S., Matthews, B. M., & Jones, C. (2005). Grey literature, institutional repositories and the organisational context. In *Seventh International Conference on Grey Literature: Open Access to Grey Resources*, 5-6 December 2005. Amsterdam: TextRelease, 2006.
- Lansing, V. C., & Carter, M. J. (2009). Does open access in ophthalmology affect how articles are subsequently cited in research? *Ophthalmology*, 116(8), 1425–31.
- Larivière, V., Zuccala, A., & Archambault, É. (2008). The declining scientific impact of theses: Implications for electronic thesis and dissertation repositories and graduate studies. *Scientometrics*, 74(1), 109–121. <http://doi.org/10.1007/s11192-008-0106-3>
- Laviña Orueta, J., & Mengual Pavón, L. (2008). *Libro Blanco de la Universidad Digital 2010*. (L. Laviña Orueta, J., & Mengual Pavón, Ed.). Barcelona, España: Ariel.
- Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature*, 411(31 May), 521.
- Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación. Boletín Oficial del Estado, núm. 131, de 2 de junio de 2011. (2011). Retrieved from <https://www.boe.es/buscar/pdf/2011/BOE-A-2011-9617-consolidado.pdf>
- Leydesdorff, L. (1998). Theories of citation? *Scientometrics*, 43(1), 5–25.
- Liau, T., Tjiek, L., & Nugraha, A. (2011). Institutional Repositories: Facilitating Structure, Collaborations, Scholarly Communications, and Institutional Visibility. In: *Digital Libraries: Methods and Applications*. InTech, Austria (pp. 163-180).
- Lin, J., & Fenner, M. (2013). The many faces of article-level metrics. *Bulletin ASIST*, (April/May).
- Lin, S.-K. (2009). Full Open Access Journals Have Increased Impact Factors. *Molecules*, 14(6), 2254–2255. article. <http://doi.org/10.3390/molecules14062254>
- Liu, Z., Min, Q., & Ji, S. (2008). A comprehensive review of research in IT adoption. In *Proceedings of the 4th International Conference on Wireless Communications, Networking and Mobile Computing* (pp. 1–5).
- Llorens Largo, F. (2011). La biblioteca universitaria como difusor de la innovación educativa: estrategia y política institucional de la Universidad de Alicante. *Arbor*, 187(Ext 3ra), 89–100. <http://doi.org/http://dx.doi.org/10.3989/arbor.2011.Extra-3n3132>
- Llorens Largo, F. (2014). Campus virtuales: De gestores de contenidos a gestores de metodologías. *RED, Revista de Educación a Distancia*, 42, 1–12.
- Llorens Largo, F., Bayona, J. J., Gómez, J., & Sanguino, F. (2010). The University of Alicante's institutional strategy to promote the open dissemination of knowledge. *Online Information Review*, 34(4), 565–582. <http://doi.org/10.1108/14684521011072981>
- López, F.-A. (2013, April 1). Visibilidad e impacto de los repositorios digitales en acceso abierto.

- López Medina, A. (31, agosto, 2006). (n.d.). Beneficios de los repositorios institucionales para la comunidad universitaria [Mensaje de Blog]. Retrieved from <http://www.madrimasd.org/blogs/openaccess/2006/08/31/39195>
- López-Yepes, J. L., Fernández Bajón, M. T., & Prat Sedeño, J. (2005). Las tesis doctorales en biblioteconomía y documentación. diagnóstico y propuesta de criterios de evaluación. *Documentación de las Ciencias de la Información*, (28), 173–187.
- Luzi, D. (2000). Trends and evolution in the development of grey literature: a review. *International Journal on Grey Literature*, 1(3), 106–117. <http://doi.org/10.1108/14666180010345537>
- Lynch, C. A. (2003). Institutional Repositories: Essential Infrastructure For Scholarship In The Digital Age. *Portal: Libraries and the Academy*, 3(2), 327–336. <http://doi.org/10.1353/pla.2003.0039>
- Mabe, M., & Price, D. (2012). Gold or green: which is the best shade of open access? Retrieved March 25, 2016, from <https://www.timeshighereducation.com/news/gold-or-green-which-is-the-best-shade-of-open-access/420454.article>
- Macha, A., & Jager, K. de. (2011). A comparative overview of the development of the institutional repositories at the University of Cape Town and at the University of Pretoria. In *14th International Symposium on Electronic Theses and Dissertations*.
- MacKenzie Owen, J. S. (1998). The expanding horizon of grey literature. In *Perspectives on the Design and Transfer of Scientific and Technical Information, Proceedings of the 3rd International Conference on Grey Literature, Luxembourg, November 13-14, 1997* (pp. 9–13). TransAtlantic, Amsterdam.
- Mahesh, G. (2012). Open access and impact factors. *Current Science*. Retrieved from <http://www.currentscience.ac.in/Volumes/103/06/0610.pdf>
- Maier, H. R. (2013). What constitutes a good literature review and why does its quality matter? *Environmental Modelling & Software*, 43, 3–4. <http://doi.org/10.1016/j.envsoft.2013.02.004>
- Markey, K., Rieh, S. Y., St. Jean, B., Yakel, E., & Yao, X. (2009). Secrets of Success: Identifying Success Factors in Institutional Repositories. In *4th International Conference on Open Repositories*. Georgia Institute of Technology.
- Marsh, C., & McLean, A. (2008). Ensuring Discovery of ETDs: The Hong Kong University of Science & Technology & ProQuest/UMI Case Study. In *ETD 2008 11th International Symposium on Electronic Theses and Dissertations, June 4-7, 2008*. Aberdeen, UK: The Robert Gordon University.
- Marsh, R. M. (2015). The role of institutional repositories in developing the communication of scholarly research. *OCLC Systems & Services*, 31(4), 163–195. <http://doi.org/10.1108/OCLC-04-2014-0022>
- Martín Rodero, H. (2014, July 9). La búsqueda bibliográfica, pilar fundamental de la Medicina Basada en la Evidencia: evaluación multivariante en las enfermedades nutricionales y metabólicas. Retrieved from <http://dspace.umh.es/handle/11000/1639>

- Martin-Martin, A., Orduna-Malea, E., Ayllon, J. M., & Lopez-Cozar, E. D. (2016). The counting house: measuring those who count. Presence of Bibliometrics, Scientometrics, Informetrics, Webometrics and Altmetrics in the Google Scholar Citations, ResearcherID, ResearchGate, Mendeley & Twitter. *EC3 Working Papers*, 21.
- McCabe, M. J., & Snyder, C. M. (2014). Identifying the effect of Open Access on citations using a panel of Science journals. *Economic Inquiry*, 52(4), 1284–1300. <http://doi.org/10.1111/ecin.12064>
- McCulloch, E. (2006). Taking stock of open access: progress and issues. *Library Review*, 55(6), 337–343. <http://doi.org/10.1108/00242530610674749>
- McDowell, C. S. (2007). Evaluating Institutional Repository Deployment in American Academe Since Early 2005: Repositories by the Numbers, Part 2. *D-Lib Magazine*, 13(910).
- McMillan, G. (2001). Do ETDs Deter Publishers?: Does Web availability count as prior publication? *College and Research Libraries News*, 62(6 (June 2001)), 620–621.
- Melero, R. (2007). Tendencias, impacto y actitudes entre los investigadores respecto al acceso abierto a las publicaciones científicas (open access). Retrieved from <http://digital.csic.es/handle/10261/1493>
- Melero, R. (2008). El paisaje de los repositorios institucionales open access en España. *BiD: Textos Universitaris de Biblioteconomia i Documentació*, (20).
- Melero, R. (2013, December 19). La diversidad de las políticas de acceso abierto. Retrieved from <http://digital.csic.es/handle/10261/88879>
- Merlo Vega, J. A., & Ferreras-Fernández, T. (2012). Cooperación para la preservación digital de contenidos locales: experiencias de las bibliotecas de las universidades públicas de Castilla y León. In *Congreso Nacional de Bibliotecas Públicas (6º. 2012. Burgos)*. Subdirección General de Coordinación Bibliotecaria.
- Merlo-vega, J. A., & Ferreras-Fernández, T. (2010). Políticas de promoción del acceso abierto en la Universidad de Salamanca. In *4as. Jornadas de OS-Repositorios*. Retrieved from <http://gredos.usal.es/xmlui/handle/10366/113357>
- Miguel Alonso, A. (2000). Aportaciones al estudio de la literatura gris universitaria. La evolución de la tesis doctoral en España. *Cuadernos de Documentación Multimedia*, (10), 26.
- Mikeal, A. (2008). ETD Management in the Texas Digital Library: lessons Learned from a Demonstrator.
- Millard, D. E., Davis, H. C., Howard, Y., McSweeney, P., Francois, S., Morris, D., ... White, S. (2010). MePrints: Building User Centred Repositories.
- Miller, P. (2000). Interoperability: What Is It and Why Should I Want It? *Ariadne*, (24).
- Ministerio de Educación. (2011). Real Decreto 99/2011, de 28 de enero, por el que se regulan las enseñanzas oficiales de doctorado. *BOE*. Madrid, Spain: Madrid, Spain: Gobierno de España.

- Moed, H. F. (2007). The effect of "open access" on citation impact: An analysis of ArXiv's condensed matter section. *Journal of the American Society for Information Science and Technology*, 58(13), 2047–2054. <http://doi.org/10.1002/asi.20663>
- Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3), 265–277. <http://doi.org/http://dx.doi.org/10.1016/j.joi.2010.01.002>
- Moralejo Sánchez, M. (2000). Las tesis doctorales de las Universidades españolas: Control bibliográfico y acceso. *Revista General de Información Y Documentación*, 10(1), 235–243.
- Morales, E. M., Gil, A. B., & García-Peñalvo, F. J. (2007). Arquitectura para la Recuperación de Objetos de Aprendizaje de Calidad en Repositorios Distribuidos. In F. G. V. & P. P. Rodríguez (Ed.), *Actas del 5º Taller en Sistemas Hipermedia Colaborativos y Adaptativos, SHCA 2007* (pp. 31–38). Zaragoza, España: SHA.
- Morales Morgado, E. M., Campos Ortuño, R. A., Yang, L. L., & Ferreras-Fernández, T. (2014). Adaptation of Descriptive Metadata for Managing Educational Resources in the GREDOS Repository. *International Journal of Knowledge Management*, 10(4), 50–72. <http://doi.org/10.4018/ijkm.2014100104>
- Morgan, P., Downing, J., Murray-Rust, P., Stewart, D., Tonge, A., Townsend, J. A., ... Rzepa, H. S. (2008). Extracting and re-using research data from chemistry e-theses: the SPECTRa-T project. In *ETD 2008 11th International Symposium on Electronic Theses and Dissertations*. 11th International Symposium on Electronic Theses and Dissertations.
- Morrison, H. (n.d.). The Imaginary Journal of Poetic Economics. Retrieved from <http://poeticeconomics.blogspot.com.es/>
- Moyle, M., Barnes, T., Brown, J., & Sadler, K. (2012, February 17). Electronic doctoral theses in the UK: a sector-wide survey into policies, practice and barriers to Open Access. UK Council for Graduate Education.
- Mueller-Langer, F., & Watt, R. (2014). The Hybrid Open Access Citation Advantage: How Many More Cites is a \$3,000 Fee Buying You? *Max Planck Institute for Innovation & ...* Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2391692
- Mukherjee, B. (2014). Green and gold open access in India. *Learned Publishing*, 27(1), 21–32. article. <http://doi.org/10.1087/20140104>
- Mulhanga, M. M., Lima, S. R., Massingue, V., & Ferreira, J. N. (2014). Expanding Scientific Knowledge Frontiers: Open Repositories in Developing Countries Supported by NRENs. In K. Rocha, A and Correia, AM and Tan, FB and Stroetmann (Ed.), *New Perspectives in Information Systems and Technologies, vol.1* (Vol. 275, pp. 127–136). http://doi.org/10.1007/978-3-319-05951-8_13
- Mulrow, C. D. (1994). Systematic reviews: Rationale for systematic reviews. *British Medical Journal*, 309, 597–599.
- Muñoz, C., García-Peñalvo, F. J., Morales, E. M., Conde, M. Á., Seoane, A. M., Gibbs, W., ... Moreira, T. (2012). Improving Learning Object Quality. *International Journal of Distance Education Technologies*, 10(4), 1–16. <http://doi.org/10.4018/jdet.2012100101>

- McGovern, N. (2009). Institutional repositories: The great debate. Topic 4: Institutional repository success is dependent upon mandates. Negative argument. *Bulletin of the American Society for Information Science & Technology*, 35(4), 27–31.
- New IEEE Standard Dictionary of Electrical, Electronics and Computer Terms*. (1994) (5th ed.). New Delhi: Galgotia Publications.
- Neylon, C. (2012). More than just access: delivering on a network-enabled literature. *PLoS Biology*, 10(10), e1001417. <http://doi.org/10.1371/journal.pbio.1001417>
- Nicholas, D., Rowlands, I., Watkinson, A., Brown, D., & Jamali, H. R. (2012). Digital repositories ten years on: What do scientific researchers think of them and how do they use them? *Learned Publishing*, 25(3), 195–206.
- Nicholson, D., & Shiri, A. (2003). Interoperability in subject searching and browsing. *OCLC Systems & Services: International Digital Library Perspectives*, 19(2), 58–61. <http://doi.org/10.1108/10650750310481775>
- Niyazov, Y., Vogel, C., Price, R., Lund, B., Judd, D., Akil, A., ... Shron, M. (2016). Open Access Meets Discoverability: Citations to Articles Posted to Academia.edu. *PLoS One*, 11(2), e0148257. <http://doi.org/10.1371/journal.pone.0148257>
- Norris, M., Oppenheim, C., & Rowland, F. (2008). Open access citation rates and developing countries. In *Open Scholarship: Authority, Community, and Sustainability in the Age of Web 2.0 - Proceedings of the 12th International Conference on Electronic Publishing, ELPUB 2008* (pp. 335–342). Department of Information Science, Loughborough University, Loughborough, Leicestershire LE11 3TU, United Kingdom.
- Norris, M., Oppenheim, C., & Rowland, F. (2008). The citation advantage of open-access articles. *Journal of the American Society for Information Science & Technology*, 59(12). <http://doi.org/10.1002/asi.20898>
- OpenAire. (n.d.). Open Access in Horizon 2020. Retrieved from <https://www.openaire.eu/open-access-in-horizon-2020>
- Orduna-Malea, E., & Delgado Lopez-Cozar, E. (2015). The dark side of open access in Google and Google Scholar: the case of Latin-American repositories. *Scientometrics*, 102(1), 829–846. article. <http://doi.org/10.1007/s11192-014-1369-5>
- Orduña-Malea, E., & Lopez-Cozar, E. D. (2014). The dark side of Open Access in Google and Google Scholar: the case of Latin-American repositories, 16. Digital Libraries. Retrieved from <http://arxiv.org/abs/1406.4331>
- Orduña-Malea, E., Martín-Martín, A., & Delgado-López-Cózar, E. (2016). The next bibliometrics: ALMetrics (Author Level Metrics) and the multiple faces of author impact. *El Profesional de La Información*, 25(3), 485. <http://doi.org/10.3145/epi.2016.may.18>
- Orera Orera, L. (2003). La edición digital de tesis doctorales: hacia la resolución de los problemas de accesibilidad. *Revista Interamericana de Bibliotecología*, 26(1).
- Otlet, P. (1934). *Traité de Documentation. Le Livre sur le Livre. Théorie et Pratique*. Brussels: Van Keerberghen.

- Paillassard, P., Schöpfel, J., & Stock, C. (2007). Dissemination and preservation of French print and electronic theses. *Grey Journal (TGJ)*, 3(2), 77–93.
- Palmer, C., Tefteau, L., & Newton, M. (2008). Strategies for institutional repository development: A case study of three evolving initiatives. *Library Trends*, 57(2), 142–167.
- Palmer, L. A. (2014). Cultivating Scholarships: The Role of Institutional Repositories in Health Sciences Libraries. *Against the Grain*, 26(2), 24–28. Park, E. G., & Richard, M. (2011). Metadata assessment in e-theses and dissertations of Canadian institutional repositories. *The Electronic Library*, 29(3), 394–407. <http://doi.org/10.1108/02640471111141124>
- Pautasso, M. (2013). Ten Simple Rules for Writing a Literature Review. *PLoS Computational Biology*, 9(7).
- Peset, F., & Ferrer, A. (2008). Implantación de la Open Archives Initiative en España. *Information Research*, 13(4).
- Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide*. *cebma.org*. <http://doi.org/10.1027/1016-9040.11.3.244>
- Phelps, S. F., & Campbell, N. (2012, September 27). Systematic Reviews in Theory and Practice for Library and Information Studies. *Library and Information Research*.
- Picarra, M., Angelaki, M., Dogan, G., Guy, M., & Artusio, C. (2015). Aligning European OA policies with the Horizon 2020 OA policy. *Insights: The UKSG Journal*, 28(3), 32–43.
- Pickton, M., & McKnight, C. (2006). Research students and the Loughborough institutional repository. *Journal of Librarianship and Information Science*, 38(4), 203–219. <http://doi.org/http://dx.doi.org/10.1177/0961000606066572>
- Pickton, M., & McKnight, C. (2007). Is there a role for research students in an institutional repository? Some repository managers' views. *Journal of Librarianship and Information Science*, 39(3), 153–161.
- Pinfield, S. (2009). Journals and repositories: an evolving relationship? *Learned Publishing*, 22(3), 165–175. <http://doi.org/10.1087/2009302>
- Pinfield, S. (2015). Making Open Access work. *Online Information Review*, 39(5), 604–636. <http://doi.org/10.1108/OIR-05-2015-0167>
- Pinfield, S., Salter, J., Bath, P. A., Hubbard, B., Millington, P., Anders, J. H. S., & Hussain, A. (2014). Open-Access Repositories Worldwide, 2005-2012: Past Growth, Current Characteristics, and Future Possibilities. *Journal of the Association for Information Science and Technology*, 65(12). <http://doi.org/10.1002/asi.23131>
- Piwovar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ*, 1, e175. <http://doi.org/10.7717/peerj.175>
- PLOS ONE. (n.d.). PLOS ONE: journal information. Retrieved March 26, 2016, from <http://journals.plos.org/plosone/s/journal-information#loc-more-information>
- Priem, J., Groth, P., & Taraborelli, D. (2012). The altmetrics collection. *PloS One*, 7(11), e48753.

<http://doi.org/10.1371/journal.pone.0048753>

- Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *arXiv12034745v1 csDL 20 Mar 2012, 1203.4745*, 1–23. Retrieved from <http://arxiv.org/abs/1203.4745>
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). *Altmetrics: a manifesto*. Retrieved from <http://altmetrics.org/manifesto/>
- Prins, A. A. M., Costas, R., van Leeuwen, T. N., & Wouters, P. F. (2016). Using Google Scholar in research evaluation of humanities and social science programs: A comparison with Web of Science data. *Research Evaluation*, rvv049. <http://doi.org/10.1093/reseval/rvv049>
- Prost, H., & Schöpfel, J. (2014). Degrees of Openness: Access Restrictions in Institutional Repositories. *D-Lib Magazine*, 20(7/8). <http://doi.org/10.1045/july2014-prost>
- Rajendiran, P., Ramesh Babu, B., & Gopalakrishnan, S. (2005). Preservation of Electronics Theses and Dissertations: A case study of SRM Institute of Science and Technology, Tamil Nadu, (India). In *ETD2005 8th International Symposium on Electronic Theses and Dissertations*. Sidney, Australia.
- Ramirez, M. L., Dalton, J. T., McMillan, G., Read, M., & Seamans, N. H. (2013). Do Open Access Electronic Theses and Dissertations Diminish Publishing Opportunities in the Social Sciences and Humanities? Findings from a 2011 Survey of Academic Publishers. *College & Research Libraries*, 74(4), 368–380.
- Ramirez, M. L., McMillan, G., Dalton, J. T., Hanlon, A., Smith, H. S., & Kern, C. (2014). Do Open Access Electronic Theses and Dissertations Diminish Publishing Opportunities in the Sciences? *College & Research Libraries*, 75(6), 808–821. <http://doi.org/10.5860/crl.75.6.808>
- Ramirez, M., & Parham, S. W. (2010). From education to preservation: Emerging institutional repository services in the scholarship lifecycle. *Proceedings of the American Society for Information Science and Technology*, 47(1), 1–2. <http://doi.org/10.1002/meet.14504701059>
- Ramírez-Montoya, M. S., & García-Peñalvo, F. J. (2015, April 11). Movimiento educativo abierto. *Virtualis*.
- Randolph, J. J. (2009). A Guide to Writing the Dissertation Literature Review. *Practical Assessment, Research & Evaluation*, 14(13).
- Rapple, C. (2011). The role of the critical review article in alleviating information overload. *Annual Reviews White Paper*.
- RCUK. (2013). *RCUK Policy on Open Access and Supporting Guidance*. Swindon. Retrieved from www.rcuk.ac.uk/documents/documents/RCUKOpenAccessPolicy.pdf
- Rehemtula, S., Rosa, M., Leitão, P., & Avilés, R. (2014). Altmetrics in institutional repositories: new perspectives for assessing research impact. *Libraries in the Digital Age (...)* Retrieved from <https://run.unl.pt/handle/10362/12398>
- Reigeluth, C. M., & Frick, T. W. (1999). Formative research: A methodology for creating and improving design theories. In C. M. Reigeluth (Ed.), *Instructional-design theories* (pp. 633--

651).

Reitz, J. M. (n.d.). Interoperability. In *ODLIS – Online Dictionary for Library and Information Science*. Retrieved from http://www.abc-clio.com/ODLIS/odlis_i.aspx

Research and Innovative Technology Administration (RITA). (n.d.). Principles for the Organization of the Transportation Research Thesaurus. Retrieved from http://ntl.bts.gov/tools/trt/trt_principles.html

Ribeiro, L., Castro, P. de, & Mennielli, M. (2016). *EUNIS - EUROCRIS joint survey on CRIS and IR: Final report*.

Riera, M., & Aibar, E. (2013). ¿Favorece la publicación en abierto el impacto de los artículos científicos? Un estudio empírico en el ámbito de la medicina intensiva. *Medicina Intensiva*, 37(4), 232–240. JOUR. <http://doi.org/10.1016/j.medin.2012.04.002>

Ríos-Hilario, A. B., Ferreras-Fernández, T., & Martín-Campo, D. (2013). Linked open bibliographic data. In *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality - TEEM '13* (pp. 333–337). New York, New York, USA: ACM Press. <http://doi.org/10.1145/2536536.2536587>

Ríos-Hilario, A. B., Ferreras-Fernández, T., & Martín-Campo, D. (2014). From Bibliographic Records to Data. *Information Resources Management Journal*, 27(3), 28–41. <http://doi.org/10.4018/irmj.2014070103>

Ríos-Hilario, A. B., Martín-Campo, D., & Ferreras-Fernández, T. (2012). Linked data y linked open data: su implantación en una biblioteca digital. El caso de Europeana. *El Profesional de la Información*, 21(3), 292–297.

Rodriguez-Marquez, M. de M., Daoutis, C. A., & Gravely, C. (2014). The training needs of PhD students: A case study from the University of Surrey. In *ETD2014*. Leicester: University of Leicester.

Rovira, C., Marcos, M.-C., & Codina, L. (2007). Repositorios de publicaciones digitales de libre acceso en Europa: análisis y valoración de la accesibilidad, posicionamiento web y calidad del código digital. *El Profesional de La Información*, 16(1), 24–38.

Ruiz-Conde, E., Calderon-Martinez, A., & Calderón-Martínez, A. (2014). University institutional repositories: competitive environment and their role as communication media of scientific knowledge. *Scientometrics*, 98(2), 1283–1299. article. <http://doi.org/10.1007/s11192-013-1159-5>

Sabharwal, S., Patel, N., & Johal, K. (2014). Open access publishing: a study of current practices in orthopaedic research. *International Orthopaedics*, 38(6), 1297–1302. article. <http://doi.org/10.1007/s00264-013-2250-5>

Sáenz, A. (2001). Leer e interpretar una revisión sistemática. *Bol Pediatr.*, 41(177), 215–21.

Sahu, A., & Goswami, N. (2015). Institutional Repository: Prospect and its global impact. *VSRD International Journal of Technical & Non-Technical Research*, 6(Special issue), 143–148.

Sale, A. (2006). The impact of mandatory policies on ETD acquisition. *D-Lib Magazine*.

- Saorín, T., & Gómez-Hernández, J.-A. (2014). Alfabetizar en tecnologías sociales para la vida diaria y el empoderamiento. *Anuario ThinkEPI*, 8, 342–348.
- Sassi, M., Marzi, C., & Pardelli, G. (2010). A terminology-based Re-definition of Grey Literature. In *Twelfth International Conference on Grey Literature ; (2010-12-06 / 2010-12-07 ; ; Praha (CZ) ;)*. Národní technická knihovna.
- Sawant, S. (2011). Institutional repositories in India: a preliminary study. *Library Hi Tech News*, 28(10), 6–10. JOUR. <http://doi.org/http://dx.doi.org/10.1108/07419051111202646>
- Schöpfel, J. (2006). Observations on the Future of Grey Literature. *The Grey Journal*, 2(2), 67–76.
- Schöpfel, J. (2012). Vers une nouvelle définition de la littérature grise. *Cahiers de La Documentation*, 66(3), 14–24.
- Schöpfel, J. (2013). Adding value to electronic theses and dissertations in institutional repositories. *D-Lib Magazine*, 19(3–4), 2.
- Schöpfel, J., Chaudiron, S. S., Jacquemin, B., Prost, H. H., Severo, M., & Thiault, F. (2014). Open access to research data in electronic theses and dissertations: an overview. *Library Hi Tech*, 32(4), 612–627. <http://doi.org/10.1108/LHT-06-2014-0058>
- Schöpfel, J., & Prost, H. (2013). Archives institutionnelles. Observations sur un nouveau mode d'information scientifique. *Revue Canadienne Des Sciences de L'information et de Bibliothéconomie*, 37(2), 122–136.
- Schöpfel, J., Prost, H., & Le Bescond, I. (2011). Open Is Not Enough: A case study on grey literature in an OAI environment. In *Thirteenth International Conference on Grey Literature* (p. 22).
- Schöpfel, J., Prost, H., Piotrowski, M., Hilf, E. R., Severiens, T., & Grabbe, P. (2014a). Restricted vs. Open Access for Electronic Theses and Dissertations – A Challenge for Public Science. In *ETD2014* (p. 9 p.).
- Schöpfel, J., Prost, H., Piotrowski, M., Hilf, E., Severiens, T., & Grabbe, P. (2014b). E-Dissertations: Access and Restrictions (EDAR) Survey 2014 Report, 1–27.
- Schöpfel, J., & Soukouya, M. (2013). Providing Access to Electronic Theses and Dissertations: A Case Study from Togo. *D-Lib Magazine*, 19(11–12). JOUR. <http://doi.org/http://dx.doi.org/10.1045/november2013-schopfel>
- Schöpfel, J., Stock, C., Farace, D. J., & Frantzen, J. (2005). Citation analysis and grey literature: Stakeholders in the grey circuit. In D. Farace (Ed.), *GL6: Work on Grey in Progress, Conference Proceedings* (pp. 55–63). inproceedings.
- Schöpfel, J., Zendulkova, D., & Fatemi, O. (2014). Electronic Theses and Dissertations in CRIS. *Procedia Computer Science*, 33, 110–117. <http://doi.org/10.1016/j.procs.2014.06.018>
- Seadle, M. (2008). The digital library in 100 years: damage control. *Library Hi Tech*, 26(1), 5–10. <http://doi.org/10.1108/07378830810857744>

- Seoane-Pardo, A. M. (2014). *Formalización de un modelo de formación online basado en el factor humano y la presencia docente mediante un lenguaje de patrón (Tesis Doctoral)*. Salamanca, España: Universidad de Salamanca. Retrieved from <http://gredos.usal.es/jspui/handle/10366/123342>
- Seoane-Pardo, A. M., & García-Peñalvo, F. J. (2014a). Patrones pedagógicos y docencia en red. In J. M. Jerónimo Montes (Ed.), *Formación en Red: Aprender con Tecnologías Digitales* (pp. 30–47). México: Universidad Nacional Autónoma de México. Facultad de Estudios Superiores Zaragoza.
- Seoane-Pardo, A. M., & García-Peñalvo, F. J. (2014b). Pedagogical Patterns and Online Teaching. In F. J. García-Peñalvo & A. M. Seoane Pardo (Ed.), *Online Tutor 2.0: Methodologies and Case Studies for Successful Learning* (pp. 298–316). Hershey, PA, USA: IGI Global.
- Setenareski, L. E., Sunye, M. S., & Shima, W. (2013). Adopting open protocols to increase the impact on digital repositories. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 8011 LNCS, pp. 408–416). SER, Universidade Federal Do Paraná, Curitiba, Brazil. http://doi.org/10.1007/978-3-642-39194-1_48#
- Shaw, C. (2013). Hundreds of open access journals accept fake science paper. *The Guardian*. Retrieved from <http://www.theguardian.com/higher-education-network/2013/oct/04/open-access-journals-fake-paper>
- Shearer, K. (2003). Institutional Repositories: Towards the Identification of Critical Success Factors. *Canadian Journal of Information and Library Science*, 27, 89–108.
- Shen, R., Goncalves, M. A., & Fox, E. A. (2013). Key Issues Regarding Digital Libraries: Evaluation and Integration. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 5(2), 1–110.
- Shotton, D. (2012). The Five Stars of Online Journal Articles - a Framework for Article Evaluation. *D-Lib Magazine*, 18(1/2). <http://doi.org/10.1045/january2012-shotton>
- Sibbald, S. L., MacGregor, J. C. D., Surmacz, M., & Wathen, C. N. (2015). Into the gray: a modified approach to citation analysis to better understand research impact. *Journal of the Medical Library Association: JMLA*, 103(1), 49–54. article. <http://doi.org/10.3163/1536-5050.103.1.010>
- Simpson, P., & Hey, J. (2006). Repositories for research: Southampton's evolving role in the knowledge cycle. *Program: Electronic Library & Information Systems*, 40(3), 224–231. article. <http://doi.org/10.1108/00330330610681303>
- Singeh, F. W., Abrizah, A., & Karim, N. H. A. (2013). What inhibits authors to self-archive in Open Access repositories? A Malaysian case. *Information Development*, 29(1), 24–35. JOUR. <http://doi.org/http://dx.doi.org/10.1177/0266666912450450>
- Sixth International Conference on Grey Literature : Work on Grey in Progress, 6-7 December 2004. (2004). In *GL6, New York (NY, US), 2004-12-06*. Amsterdam, Netherlands: TextRelease.

- Snijder, R. (2010). The profits of free books: an experiment to measure the impact of open access publishing. *Learned Publishing*, 23(4), 293–301. <http://doi.org/10.1087/20100403>
- Solomona, D. J., Laaksob, M., & Bjork, B.-C. (2013). A longitudinal comparison of citation rates and growth among open access journals. *Journal of Informetrics*, 7(3), 642–650. <http://doi.org/http://dx.doi.org/10.1016/j.joi.2013.03.008>
- Søndergaard, T. F., Andersen, J., & Hjørland, B. (2003). Documents and the communication of scientific and scholarly information. *Journal of Documentation*, 59(3), 278–320. <http://doi.org/10.1108/00220410310472509>
- Sorli Rojo, Á., & Merlo-Vega, J. A. (2002). Bases de datos y recursos en internet de tesis doctorales. *Revista Española de Documentación Científicas*, 25(2).
- SPARC Europe. (2015). The Open Access Citation Advantage: List of studies and results to date. Retrieved from http://sparceurope.org/oaca_table/
- SPARC, PLOS, & OASPA. (2012). HowOpenIsIt? Retrieved from <https://www.plos.org/how-open-is-it>
- Stake, R. E. (1995). *Investigación con estudio de casos*. Madrid: Morata.
- Stanton, K. V., & Liew, C. L. (2011). Open access theses in institutional repositories: an exploratory study of the perceptions of doctoral students. *Information Research: An International Electronic Journal*, 17(1).
- Sterman, L. (2014). Institutional Repositories: An Analysis of Trends and a Proposed Collaborative Future. *College & Undergraduate Libraries*, 21(3–4), 360–376. <http://doi.org/10.1080/10691316.2014.943919>
- Stone, S. M., & Lowe, M. S. (2014). Who is Citing Undergraduate Theses in Institutional Digital Repositories? Implications for Scholarship and Information Literacy. *College & Undergraduate Libraries*, 21(3/4), 345–359. <http://doi.org/10.1080/10691316.2014.929065>
- Suber, P. (2004). A Very Brief Introduction to Open Access. *Living Reviews*. Retrieved from <http://www.livingreviews.org/open-access.html>
- Suber, P. (2012a). Ensuring open access for publicly funded research. *BMJ (Clinical Research Ed.)*, 345, e5184. <http://doi.org/10.1136/bmj.e5184>
- Suber, P. (2012b). Open Access. MIT Press. Retrieved from <http://dash.harvard.edu/handle/1/10752204>
- Suber, P. (2015). *Acceso abierto* (1ª). Toluca, Estado de México: Universidad Autónoma del Estado de México.
- Subirats, I., Onyanha, I., Salokhe, G., Kaloyanova, S., Anibaldi, S., & Keizer, J. (2013). Towards an architecture for open archive networks in agricultural sciences and technology. *Online Information Review*.
- Sud, P., & Thelwall, M. (2014). Evaluating altmetrics. *SCIENTOMETRICS*, 98(2), 1131–1143. article. <http://doi.org/10.1007/s11192-013-1117-2>

- Sutton, S. C. (2013). Open access, publisher embargoes, and the voluntary nature of scholarship: An analysis. *College & Research Libraries News*, 74(9), 468–472.
- Swan, A. (2007). The business of digital repositories. In K. Weenink, L. Waaije, & K. van Godtsenhoven (Eds.), *A DRIVER'S guide to European repositories* (pp. 15–47). Amsterdam, Netherlands: Amsterdam University Press.
- Swan, A. (2010). The Open Access citation advantage: Studies and results to date. *Research Report by Alma Swan*, 3, 17.
- Swan, A., & Awre, C. (2006). *Linking UK Repositories: Technical and Organisational Models to Support User-Oriented Services Across Institutional and Other Digital Repositories*. London: Jisc.
- Swan, A., & Carr, L. (2008). Institutions, Their Repositories and the Web. *Serials Review*, 34(1), 31–35. <http://doi.org/10.1016/j.serrev.2007.12.006>
- Swan, A., Gargouri, Y., Hunt, M., & Harnad, S. (2015). *Open Access Policy: Numbers, Analysis, Effectiveness* (Digital Libraries). Retrieved from <http://arxiv.org/abs/1504.02261>
- Syfret, M. R. H. (1948). The Origins of the Royal Society. *Notes and Records of the Royal Society*, 5(2), 75–137.
- Templier, M., & Paré, G. (2015). A Framework for Guiding and Evaluating Literature Reviews. *Communications of the Association for Information Systems* 37(1-6).
- Terras, M. (2012). The impact of social media on the dissemination of research: results of an experiment. *Journal of Digital Humanities*, 1(3).
- The Writing Lab, The OWL at Purdue, & Purdue University. (2016). APA Style. Retrieved from <https://owl.english.purdue.edu/owl/section/2/10/>
- Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do Altmetrics Work? Twitter and Ten Other Social Web Services. *PLoS ONE*, 8(5), e64841. <http://doi.org/10.1371/journal.pone.0064841>
- Therón, R., & Fontanillo, L. (2015). Diachronic-information visualization in historical dictionaries. *Information Visualization*, 14, 11–136.
- Therón, R., Seguí, C., Cruz, L., & Vaquero, M. (2014). Highly interactive and natural user interfaces: enabling visual analysis in historical lexicography. In *DATECH '14 Proceedings of the First International Conference on Digital Access to Textual Cultural Heritage* (pp. 153–158). New York, NY, USA: ACM.
- Therón, R., & Wandl-Vogt, E. (2014). The fun of exploration: How to access a non-standard language corpus visually. In A. Hautli-Janisz, V. Lyding, & C. Rohrdantz (Eds.), *VisLR: Visualization as Added Value in the Development, Use and Evaluation of Language Resources* (pp. 9–12).
- Thibodeau, K. (2007). If you build it, will it fly? Criteria for success in a digital repository. *Journal of Digital Information*, 8(2).

- Third International Conference on Grey Literature : Perspectives on the Design and Transfer of Scientific and Technical Information, 13-14 November 1997. (1997). In *GL3, Luxembourg (LU), 1997-11-13*. Amsterdam, Netherlands: GreyNet.
- Thorin, S. (2006). Global changes in scholarly communication. *eLearning and Digital Publishing*. New York: Springer Verlag.
- Timeline of the open access movement. (n.d.). Retrieved from <http://oad.simmons.edu/oadwiki/Timeline>
- Torres-Salinas, D., Ruiz-Pérez, R., & Delgado-López-Cózar, E. (2009). Google Scholar como herramienta para la evaluación científica. *El Profesional de La Información*, 18(5), 501–510. <http://doi.org/DOI: 10.3145/epi.2009.sep.03>
- Tripathi, M., & Jeevan, V. K. J. (2011). An Evaluation of Digital Libraries and Institutional Repositories in India. *The Journal of Academic Librarianship*, 37(6), 543–545. <http://doi.org/10.1016/j.acalib.2011.08.012>
- Troman, A., Jacobs, N., & Copeland, S. (2007). A new electronic service for UK theses: access transformed by EThOS. *Interlending & Document Supply*, 35(3), 157–163. <http://doi.org/10.1108/02641610710780836>
- Truth, F. (2012). Pay big to publish fast: Academic journal rackets. *Journal for Critical Education Policy Studies*, 10(2), 54–105.
- Ubogu, F., & Sayed, Y. (2008). Management of Research Data in ETD Systems. In *ETD 2008 11th International Symposium on Electronic Theses and Dissertations*. Aberdeen, UK.: The Robert Gordon University.
- Unzué, M., & Freibrun, N. (2015). Políticas Públicas de Acceso abierto y democratización del conocimiento en la Universidad Argentina. *AVATARES de la Comunicación y la Cultura*, 9. Retrieved from <http://ppct.caicyt.gov.ar/index.php/avatares/article/view/6374>.
- Valverde Berrocoso, J. (2013). *El acceso abierto al conocimiento científico*. (S. i E. E. C. Universitat de Barcelona. GRC Esbrina, Ed.). Red Universitaria de Investigación Innovación Educativa (REUNI+D). Retrieved from <http://diposit.ub.edu/dspace/handle/2445/36335>
- Vanclay, J. K. (2013). Factors affecting citation rates in environmental science. *Journal of Informetrics*, 7(2), 265–271. article. <http://doi.org/10.1016/j.joi.2012.11.009>
- Varey, R. J., Wood-Harper, T., & Wood, B. (2002). A theoretical review of management and information systems using a critical communications theory. *Journal of Information Technology*, 17(4), 229–239.
- Varshney, L. R. (2012). The Google effect in doctoral theses. *Scientometrics*, 92(3), 785–793. article. <http://doi.org/10.1007/s11192-012-0654-4>
- Vaughn, J. (2013, January 1). The future of scholarly communication: US efforts to bring warring factions to common purpose in support of scholarship. *Information Services & Use*. IOS Press. <http://doi.org/10.3233/ISU-130689>
- Veiga de Cabo, J., & Martín-Rodero, H. (2011). Acceso Abierto: nuevos modelos de edición

científica en entornos web 2.0. *Salud Colectiva*, 7, S19–S27.

- Vijayakumar, J. K., & Vijayakumar, M. (2007). Importance of doctoral theses and its access: a literature analysis. *Grey Journal (TGJ)*, 3(2), 67–75.
- Vincent-Lamarre, P., Boivin, J., Gargouri, Y., Larivière, V., & Harnad, S. (2014). Estimating Open Access Mandate Effectiveness: I. The MELIBEA Score, 22. *Digital Libraries*.
- Vincent-Lamarre, P., Boivin, J., Gargouri, Y., Larivière, V., & Harnad, S. (2014, July 10). The effect of Open Access mandate strength on deposit rate and latency. UQAM Institute of Cognitive Sciences.
- Waddington, S., Green, R., & Awre, C. (2012). CLIF: Moving repositories upstream in the content lifecycle. *Journal of Digital Information*, 13(1).
- Wagner, A. Ben. (2010). Open access citation advantage: An annotated bibliography. *Issues in Science and Technology Librarianship*, 60. article.
- Walker, E. P. (2011). What We Can Learn from ETDs: Using ProQuest Dissertations & Theses as a Dataset. In *USETDA 2011: The Magic of ETDs... Where Creative Minds Meet*. Orlando, Florida.
- Walsh, J. P., Kucker, S., Maloney, N. G., & Gabbay, S. (2000). Connecting minds: Computer-mediated communication and scientific work. *Journal of the American Society for Information Science*, 51(14), 1295–1305. [http://doi.org/10.1002/1097-4571\(2000\)9999:9999<::AID-AS11046>3.0.CO;2-W](http://doi.org/10.1002/1097-4571(2000)9999:9999<::AID-AS11046>3.0.CO;2-W)
- Wanden-Berghe, C., & Sanz-Valero, J. (2014). Revisiones sistemáticas sobre las funciones de los ácidos grasos poliinsaturados omega-3 en la salud y la enfermedad. In *Gil Hernández A., Serra Majem L. Libro Blanco de los Omega-3 (eBook online)*. (2ª, pp. 73–79). Madrid: Panamericana.
- Warner, S. (2003). E-prints and the Open Archives Initiative. *Library Hi Tech*, 21(2), 151–158. <http://doi.org/10.1108/07378830310479794>
- Wellcome Trust. (2012). Wellcome Trust strengthens its open access policy. Retrieved from www.wellcome.ac.uk/News/Media-office/Press-releases/2012/WTVM055745.htm#
- Wellcome Trust. (2014). Charity open access fund. Retrieved from www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Open-access/Charity-open-access-fund/index.htm#
- Westell, M. (2006). Institutional repositories: proposed indicators of success. *Library Hi Tech*, 24(2), 211–226. <http://doi.org/10.1108/07378830610669583>
- White, W. (2007). Opening access and closing risk: delivering the mandate for e-theses deposit. In *ETD 2007 added values to e-theses: 10th International Symposium on Electronic Theses and Dissertations* (p. 6pp.). Uppsala, Sweden.
- Wilkinson, D., Sud, P., & Thelwall, M. (2014). Substance without citation: evaluating the online impact of grey literature. *Scientometrics* 98(2), 797-806.
- Williams, M. D., Dwivedi, Y. K., Lal, B., & Schwarz, A. (2009). Contemporary trends and issues in

- IT Adoption and diffusion research. *Journal of Information Technology*, 24(1), 1–10.
<http://doi.org/http://dx.doi.org/10.1057/jit.2008.30>
- Willinsky, J. (2005). *The Access Principle: The Case for Open Access to Research and Scholarship*. Cambridge, MA: The MIT Press.
- Willmott, M., Dunn, K., & Duranceau, E. (2012). The Accessibility Quotient: A New Measure of Open Access. *Journal of Librarianship and Scholarly Communication*, 1(1), eP1025.
<http://doi.org/10.7710/2162-3309.1025>
- Withey, L., Cohn, S., Faran, E., Jensen, M., Kiely, G., Underwood, W., ... Keane, K. (2011). Sustaining Scholarly Publishing: New Business Models for University Presses: A Report of the AAUP Task Force on Economic Models for Scholarly Publishing. *Journal of Scholarly Publishing*, 42(4), 397–441.
- Wouters, P., & Costas, R. (2012). Users , narcissism and control – tracking the impact of scholarly publications in the 21 st century. *Image Rochester NY*, 50 pages.
- Xia, J. (2012). Diffusionism and open access. *JOURNAL OF DOCUMENTATION*, 68(1), 72–99. article. <http://doi.org/10.1108/00220411211200338>
- Xia, J., Gilchrist, S. B., Smith, N. X. P., Kingery, J. A., Radecki, J. R., Wilhelm, M. L., ... Mahn, A. J. (2012). A Review of Open Access Self-Archiving Mandate Policies. *Portal: Libraries and the Academy*, 12(1), 85–102. JOUR. <http://doi.org/10.1353/pla.2012.0000>
- Xia, J., Harmon, J. L., Connolly, K. G., Donnelly, R. M., Anderson, M. R., & Howard, H. A. (2015). Who publishes in “predatory” journals? *Journal of the Association for Information Science and Technology*, 66(7), 1406–1417. article. <http://doi.org/10.1002/asi.23265>
- Xia, J., Myers, R. L., & Wilhoite, S. K. (2011). Multiple open access availability and citation impact. *Journal of Information Science*, 37(1), 19–28. article. <http://doi.org/10.1177/0165551510389358>
- Xia, J., & Nakanishi, K. (2012). Self-selection and the citation advantage of open access articles. *Online Information Review*, 36(1), 40–51. article. <http://doi.org/10.1108/14684521211206962>
- Xia, J., & Sun, L. (2007). Assessment of Self-Archiving in Institutional Repositories: Depositorship and Full-Text Availability. *Serials Review*, 33(1), 14–21. <http://doi.org/10.1080/00987913.2007.10765087>
- Xu, J., Kang, Q., & Song, Z. (2015). The current state of systematic reviews in library and information studies. *Library & Information Science Research*, 37(4), 296–310. <http://doi.org/10.1016/j.lisr.2015.11.003>
- Yang, H., & Tate, M. (2009). Where are we at with Cloud Computing?: A Descriptive Literature Review. In *20th Australasian Conference on Information Systems* (p. 13).
- Yin, R. (1993). *Applications of case study research*. Beverly Hills, CA: Sage Publishing.
- Yin, R. K. (1994). *Case Study Research. Design and Methods*. London: SAGE.

- Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of “alternative metrics” in scientific publications. *Scientometrics*. Retrieved from <http://arxiv.org/ftp/arxiv/papers/1404/1404.1301.pdf>
- Zhang, Y. (2006). The Effect of Open Access on Citation Impact: A Comparison Study Based on Web Citation Analysis. *Libri*, 56(3), 145–156. JOUR. <http://doi.org/10.1515/LIBR.2006.145>
- Zhao, R., & Wu, S. (2014). Study on themes and authors' influence of open access in China. *Scientometrics*, 101(2), 1165–1177.
- Zhao, Y., Niu, Z., Cao, Y., & Dai, L. (2010). Research on Evaluation of Digital Library. In *2010 International Conference on Data Storage and Data Engineering* (pp. 123–126). <http://doi.org/10.1109/DSDE.2010.20>

Anexo VI.1. Índice de tablas

	Página
Tabla 1. Estructura de la tesis	53
Tabla 2. Las 10 reglas simples para escribir una revisión de la literatura.....	85
Tabla 3. Tipos de revisiones de la literatura y características según Grant y Booth (2009).....	88
Tabla 4. Tipos de revisiones de la literatura y características según Petticrew y Roberts (2006).....	89
Tabla 5. Modelo PICOC.....	96
Tabla 6. Otros recursos utilizados para localizar la literatura	102
Tabla 7. Resultados obtenidos de WOS a la primera cuestión	103
Tabla 8. Resultados obtenidos en Scopus a la primera cuestión	103
Tabla 9. Resultados obtenidos en LISTA a la primera cuestión.....	104
Tabla 10. Resultados obtenidos en LISA a la primera cuestión.....	104
Tabla 11. Búsqueda y resultados en WOS de la segunda cuestión.....	105
Tabla 12. Búsqueda y resultados en SCOPUS de la segunda cuestión.....	105
Tabla 13. Búsqueda y resultados en LISTA de la segunda cuestión.....	106
Tabla 14. Búsqueda y resultados en LISA de la segunda cuestión.....	106
Tabla 15. Búsqueda y resultados en WOS de la tercera cuestión	107
Tabla 16. Búsqueda y resultados en SCOPUS de la tercera cuestión.....	107
Tabla 17. Búsqueda y resultados en LISTA de la tercera cuestión.....	107
Tabla 18. Búsqueda y resultados en LISA de la tercera cuestión.....	108
Tabla 19. Búsqueda y resultados en WOS de la cuarta cuestión.....	108
Tabla 20. Búsqueda y resultados en SCOPUS de la cuarta cuestión.....	109
Tabla 21. Búsqueda y resultados en LISTA de la cuarta cuestión.....	109
Tabla 22. Búsqueda y resultados en LISA de la cuarta cuestión.....	109

Tabla 23. Búsqueda y resultados en WOS de la quinta cuestión.....	110
Tabla 24. Búsqueda y resultados en SCOPUS de la quinta cuestión.....	110
Tabla 25. Búsqueda y resultados en LISTA de la quinta cuestión.....	111
Tabla 26. Búsqueda y resultados en LISA de la quinta cuestión.....	111
Tabla 27. Tabla Excel de las Búsquedas en Google Scholar.....	114
Tabla 28. Formulario de extracción de datos.....	117
Tabla 29. Revisiones y bibliografías.....	121
Tabla 30. Tópicos y temas de interoperabilidad	196
Tabla 31. Estudios sobre la ventaja de citación del acceso abierto.....	252-260
Tabla 32. Datos TESEO. Nº de tesis de la U. Salamanca por cursos.....	303
Tabla 33. Clasificación de las tesis doctorales por categorías.....	303
Tabla 34. Tesis de la Universidad de Salamanca.....	307
Tabla 35. Tesis doctorales citadas por años y nº total de citas por años.....	319
Tabla 36. Distribución de las tesis doctorales por áreas de conocimiento.....	321
Tabla 37. Distribución de las tesis doctorales por años.....	322
Tabla 38. Características de los indicadores de visibilidad e impacto según las áreas	325
Tabla 39. Características de los indicadores de visibilidad e impacto según los años.....	326
Tabla 40. Correlaciones establecidas entre las variables.....	327
Tabla 41. Distribución de las citas por área de conocimiento.....	328
Tabla 42. Distribución de las citas por años.....	328

Anexo VI.2. Índice de figuras

	Páginas
Figura 1. Estructura de la tesis. Fuente: elaboración propia realizada con <i>Cmap Tools</i>	59
Figura 2. Modificado de Maier (2013). Representación conceptual del impacto de la revisión de la literatura sobre Objetivos, Métodos, Resultados y Conclusiones.....	84
Figura 3. Cuestiones previas a la realización de una revisión sistemática. Fuente: elaboración propia.....	92
Figura 4. Diagrama de flujo de la selección de estudios para la inclusión en la revisión sistemática	116
Figura 5. Open Access Impact. Fuente: Google Trends	120
Figura 6. Tipología de la literatura obtenida en la revisión.....	122
Figura 7. Títulos de revistas más frecuentes y nº de referencias.....	123
Figura 8. Tipología de la literatura seleccionada.....	124
Figura 9. Ciclo de vida de la investigación (JIRCS).....	136
Figura 10. Organismos financiadores de investigación y mandatos (312-03-2016).....	155
Figura 11. Gráfico de Políticas adoptadas por trimestre (31-03-2016).....	155
Figura 12. Políticas según el tipo de institución.....	156
Figura 13. Crecimiento mundial de Repositorios.	171
Figura 14. Crecimiento mundial de Repositorios Institucionales (RIs).	171
Figura 15. Tipos de repositorios de acceso abierto en el mundo.....	176
Figura 16. Áreas de interoperabilidad del repositorio Gredos.	193
Figura 17. Contenido de la comunidad Biblioteca Digital de Gredos	205
Figura 18. Contenido de la comunidad Repositorio Científico de Gredos.....	206
Figura 19. Contenido de la comunidad Repositorio Docente de Gredos	207
Figura 20. Contenido del Archivo Institucional de Gredos.....	207
	393

Figura 21. Datos de Gredos proporcionados por OpenAire	209
Figura 22. Tipos de contenidos de Gredos. Datos extraídos de OpenAire.....	209
Figura 23. Crecimiento mundial de tesis electrónicas en los repositorios institucionales.....	230
Figura 24. Servicios de infraestructura del repositorio.....	248
Figura 25. Nube de etiquetas con términos usados por los investigadores con un perfil público en <i>Google Scholar Citations</i>	271
Figura 26. Herramientas de las altmétricas.....	282
Figura 27. Comunidad Tesis doctorales del repositorio Gredos y sus colecciones.....	288
Figura 28. Tamaño de la colección de Tesis doctorales del repositorio Gredos (2016).....	289
Figura 29. Diseminación del contenido del repositorio Gredos.....	296
Figura 30. Informe mensual de Gredos de Google Analytics.....	298
Figura 31. Base de datos TESEO. Pantalla de consulta.....	301
Figura 32. Porcentaje de tesis por áreas.....	303
Figura 33. Tesis Doctorales en la Universidad de Salamanca (1977-2014).....	304
Figura 34. Tesis Doctorales en la Universidad de Salamanca (1977-2014).....	305
Figura 35. Ejemplo de la base de datos construida en Excel.....	308
Figura 36. Búsqueda de tesis de la Universidad de Salamanca de 2006-2007.....	310
Figura 37 Índice de referencias citadas de WOS.....	319
Figura 38. Tesis OA USAL.....	321
Figura 39. Distribución de las visitas, descargas y citas por áreas y años.....	322
Figura 40. Representación de valores medios y error estándar de visitas por áreas y años...	323
Figura 41. Representación de valores medios y error estándar de descargas por áreas y años.	323
Figura 42. Representación de valores medios y error estándar de citas por áreas y años.....	324
Figura 43. Porcentaje de citas.....	327

Anexo VI.3. Revisión sistemática de la literatura

El anexo VI.3. contiene los *datasets* resultado de la revisión sistemática de la literatura llevada a cabo en esta investigación. Además de estar incluidos en la carpeta Anexos de esta tesis, una vez leída esta se publicarán en acceso abierto en el repositorio Gredos y como Open Data en el repositorio Zenodo.

El contenido de este anexo es el siguiente:

1. Tabla 3.1. Archivo completo con los criterios de selección aplicados en la revisión sistemática
2. Tabla 3.2. Archivo completo de las búsquedas en las bases de datos – 552 referencias
3. Tabla 3.3. Referencias seleccionadas de las bases de datos WOS, SCOPUS, LISTA y LISA – 257 referencias
4. Tabla 3.4. Revisión manual – 104 referencias
5. Tabla 3.5. Referencias seleccionadas de Google Scholar – 182 referencias
6. Tabla 3.6. Revisiones y bibliografías

Anexo VI.4. Base de datos de las tesis doctorales analizadas

El anexo VI.4. contiene las bases de datos elaboradas en Excel con los datos sobre tesis doctorales de la Universidad y los correspondientes gráficos. Al igual que el anexo VI.3, este también estará además de estar incluidos en la carpeta Anexos de esta tesis, una vez leída esta se publicarán en acceso abierto en el repositorio Gredos y como Open Data en el repositorio Zenodo.

Contiene los siguientes datasets:

1. Base de datos extraída de Gredos y TESEO (fecha de actualización: 13-04-2016)
2. Datos de la Tesis USAL en TESEO
3. Base de datos definitiva utilizada en el análisis (fecha de actualización:26-04-2016)

