



Revista española de documentación científica

vol. 39, n. 1 (2016)

Estudios

Los sectores institucionales en la producción científica española de difusión internacional

Borja González-Albo, Javier Aparicio, Luz Moreno, María Bordons

Análisis de la producción científica hispana en mejora continua: 1990-2011

Lidia Sánchez, Beatriz Blanco

Propuesta de un Centro de Archivo del Feminismo tras el análisis de los fondos documentales del Movimiento Feminista

María Adelina Codina-Canet, Rosa San Segundo Manuel

Factores determinantes en la producción científica de los grupos de investigación en Colombia

Gladys Rueda-Barrios, Manuel Rodenes-Adam

Ranking de las bibliotecas universitarias españolas en la gestión del personal

José Simón-Martín, Clara Simón-Blas, Alicia Arias-Coello

Evaluación de la accesibilidad web de los portales del Estado en Perú

Carlos Antonio Sam-Anlas, Yudayly Stable-Rodríguez

Aplicación del modelo de datos RDF en las colecciones digitales de bibliotecas, archivos y museos de España

Andreu Sulé, Miquel Centelles, Jorge Franganillo, Jesús Gascón

Notas y experiencias

“Desktop Scientometrics”: una metodología para el uso de datos procedentes de WoS mediante el programa estadístico R

Preiddy Efrain-García, Carlos García-Zorita

Sobre la elaboración de una taxonomía de Historia medieval portuguesa: problemas y desafíos

Filipa Medeiros, Hermínia Vasconcelos Vilar, José Antonio Moreira González

Crítica de libros

Gephi Cookbook. Devangana Khokhar. Birmingham, Packt publishing, 2015. 274 pags. ISBN 978-1-78398-740-5

Carlos Benito Amat

Noticias

El Centro de Ciencias Humanas y Sociales del CSIC galardonado con el Premio MEDES 2015 en la categoría de mejor institución por su labor de coordinación y mantenimiento del Sistema LATINDEX en España

Teresa Abejón Peña



ESTUDIOS / RESEARCH STUDIES

Los sectores institucionales en la producción científica española de difusión internacional

Borja González-Albo*, Javier Aparicio*, Luz Moreno**, María Bordons**

*Unidad Transversal de Apoyo a la Investigación (UTAI), Centro de Ciencias Humanas y Sociales (CCHS), Consejo Superior de Investigaciones Científicas (CSIC).

**Grupo de Análisis Cuantitativo en Ciencia y Tecnología (ACUTE), Instituto de Filosofía (IFS), Centro de Ciencias Humanas y Sociales (CCHS), Consejo Superior de Investigaciones Científicas (CSIC).

Correo -e: borja.gonzalezalbo@cchs.csic.es

Recibido: 16-02-2015; 2ª versión: 26-06-2015; Aceptado: 15-07-2015.

Cómo citar este artículo/Citation: González-Albo, B.; Aparicio, J.; Moreno, L.; Bordons, M. (2016). Los sectores institucionales en la producción científica española de difusión internacional. *Revista Española de Documentación Científica*, 39(1): e115. doi: <http://dx.doi.org/10.3989/redc.2016.1.1260>

Resumen: Los sectores institucionales, como agrupaciones de centros con características similares en su actividad y orientación a la investigación, constituyen una interesante unidad de análisis que puede servir además como marco de referencia en los estudios de I+D. El análisis de la producción científica de España (2000–2011) recogida en la base de datos Web of Science permite observar el predominio del sector universitario (66%), seguido de los Organismos Públicos de Investigación (OPI) (23%) y del sector sanitario (18%). Las entidades sin ánimo de lucro (ESAL) (10%), que incluyen tanto las Entidades Tradicionales sin Ánimo de Lucro (ETSAL) como las nuevas Estructuras de Investigación sin Ánimo de Lucro (EISAL), muestran junto con los OPI los valores más altos de impacto, tanto a través de indicadores basados en el factor de impacto de sus revistas de publicación – posición normalizada, porcentaje de artículos en primer cuartil – como de las citas recibidas por sus artículos. La Universidad presenta el perfil de producción más distribuido entre las distintas áreas temáticas, y es el único sector especializado en determinados campos como Ciencias Sociales, Humanidades o Matemáticas, mientras que el Sector Sanitario muestra una alta especialización en Medicina Clínica. A lo largo del periodo los distintos sectores tienden a aumentar su producción, obtener indicadores de impacto más elevados e incrementar su investigación en colaboración e internacionalización. Destaca el auge experimentado por las entidades sin ánimo de lucro, y en particular por las EISAL, que incrementan de forma muy llamativa su producción y realizan investigación de alto impacto relativo, sobre todo en el área biomédica. No se observa relación entre las áreas de especialización de los sectores y aquellas en las que obtienen el mayor impacto.

Palabras clave: Sectores institucionales; España; indicadores bibliométricos; Web of Science.

Institutional sectors in the international scientific production from Spain

Abstract: Institutional sectors —groups of centers with similar characteristics and types of R&D— constitute an interesting unit of analysis that can also serve as a useful framework for benchmarking purposes in R&D studies. The analysis of Spanish scientific output included in the Web of Science database (2000–2011) shows the predominance of Universities (66%), followed by the Public Research Organisations (OPI) (23%) and the Health Sector (18%). Non-profit organizations (ESAL) (10%), which include both Traditional Entities with non-Profit Aims (ETSAL) and new Research Structures with non-Profit Aims (EISAL), show the highest impact values, regardless of whether indicators based on impact factors (normalised position, percentage of articles in the first quartile) or on citations are used. Universities present a more evenly distributed thematic profile and it is the only sector specialized in fields such as Social Sciences, Humanities or Mathematics, while the Health Sector is highly specialized in Clinical Medicine. During the period studied, the different sectors tended to increase their scientific production, to publish in higher impact factor journals and to intensify the level of collaboration and internationalization of their publications. Especially noteworthy is the strong rise of non-profit organizations, in particular the case of EISAL, which conduct high impact research mainly in Biomedicine. No relationship has been found between the areas in which the sectors are specialized and those in which they obtain the highest impact.

Keywords: Institutional sectors; Spain; bibliometric indicators; Web of Science.

Copyright: © 2016 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0.

1. INTRODUCCIÓN Y OBJETIVOS

Los análisis bibliométricos centrados en la actividad científica de los sectores institucionales de un país no son muy numerosos porque su delimitación no es sencilla y constituyen grupos de instituciones con cierto grado de heterogeneidad. Sin embargo, la actividad de estas agrupaciones se describe someramente en muchos estudios que analizan la actividad científica de un país (Sort García y Rodríguez Méndez, 2012; Moya y otros, 2013) o como un primer acercamiento a otras unidades de análisis como centros o instituciones concretas para contextualizarlos (Bordons y otros, 2014). Existen en la literatura algunos estudios centrados en el papel de los distintos sectores institucionales en la actividad científica de un país, bien de forma global (Godin y Gingras, 2000; De Filippo y otros, 2013) o en determinadas áreas temáticas (Lander, 2013), pero son más frecuentes los trabajos que analizan la actividad de un determinado sector como las empresas (Moya-Anegón y otros, 2014; Chang, 2014) o las universidades (Huang y otros, 2006; Bordons y otros, 2010; Grupo Scimago, 2007). Especial mención hay que hacer de los estudios sobre la colaboración inter-sectorial, y más concretamente la colaboración entre la universidad y la empresa, que ha despertado un gran interés en la literatura por su importancia para facilitar la transferencia de conocimiento entre el sector público y privado (Abramo y otros, 2009).

El Manual de Frascati, que aporta definiciones y categorías de actividades a tener en cuenta para la obtención de estadísticas de I+D, señala la importancia de la clasificación institucional de los esfuerzos nacionales de I+D por sectores, ya que los sectores difieren en sus características y orientación a la investigación. La sectorialización institucional facilita la recogida de datos de I+D, pero también su análisis e interpretación (OCDE, 2003). De hecho, los indicadores de actividad científica de un país a nivel sector proporcionan una información interesante y rica sobre algunos aspectos de la estructura y dinámica de la investigación en el contexto nacional. Tal y como apunta Lander (2013), estudiar el papel de los diferentes sectores institucionales y las relaciones que se establecen entre ellos, influidas por sus distintas motivaciones y formas de colaborar, es clave para establecer un modelo realista y preciso de los procesos de I+D. Por otro lado, los indicadores a nivel de sectores institucionales permiten establecer puntos de referencia para la posterior comparación de las distintas entidades de investigación que se encuentran incluidas en cada una de las agrupaciones. Finalmente, es posible ana-

lizar el rendimiento de los distintos sectores en función del esfuerzo y dedicación que invierten en tareas de investigación si se cuenta con datos de input o estadísticas de I+D recogidos con el mismo criterio.

En este contexto, el presente trabajo contempla los siguientes objetivos:

1. Describir la actividad investigadora de los sectores institucionales españoles en el periodo 2000–2011 a partir de su producción científica incluida en la base de datos Web of Science (WoS), centrándose en las siguientes cuestiones:
 - ¿Cuál es su contribución al conjunto de la producción científica española?
 - ¿Existen diferencias entre sectores institucionales en su especialización temática y en su impacto científico? ¿Obtienen los sectores su mayor impacto en sus áreas de especialización?
 - ¿Se observan diferencias en el patrón de colaboración de los distintos sectores?
2. Analizar la evolución de los distintos sectores institucionales, comparando su actividad en el primer y último bienio del periodo analizado:
 - ¿Se aprecian cambios en su contribución a la producción total del país?
 - ¿Se mantiene estable su especialización temática? ¿Qué sectores muestran una evolución más positiva en su impacto?
 - ¿Cuál es la evolución de los distintos sectores en lo que se refiere a su actividad en colaboración?

2. REVISIÓN BIBLIOGRÁFICA Y JUSTIFICACIÓN

Datos de producción científica desagregados por sectores institucionales se encuentran en los informes que analizan la actividad científica y tecnológica de países como Holanda (Tijssen y otros, 2010), Japón (Science and Technology Indicator Project Team, 2004), Finlandia (Lehvo y Nuutinen, 2006) o Estados Unidos (National Science Board, 2014). En el caso de España, cabe destacar la publicación "Indicadores bibliométricos de la actividad científica española 2010", editada por FECYT, que incluye un apartado con un análisis descriptivo de la actividad, colaboración y el impacto de los distintos sectores institucionales del país (Moya-Anegón y otros, 2013). Así mismo, el grupo ACUTE incluye de forma habitual resultados acerca de sectores en los informes realizados sobre el CSIC (Bordons y otros, 2014) o la Comuni-

dad de Madrid (Gómez y otros, 2009). A nivel regional se pueden citar los estudios sobre Cataluña (Rovira y otros, 2007) o Andalucía (Moya-Anegón y otros, 2008) que presentan gran parte de sus resultados por sectores institucionales.

Sin embargo, los artículos científicos centrados en el estudio comparativo de la actividad de uno o más países por sectores institucionales no son muy numerosos. Se puede señalar el estudio de Katz y Hicks (1996), que analiza la actividad científica británica utilizando como base la distribución en sectores institucionales, y destaca el importante papel jugado por la Universidad. De un modo similar, Godin y otros (2002) realizaron un estudio de la producción científica canadiense por sectores institucionales, profundizando en la aportación de la Universidad, la relación de ésta con la industria y la necesidad de una política activa para fomentar dicha colaboración. Otros estudios analizan el papel que juegan los diferentes sectores institucionales en la investigación sobre temas específicos (Lander, 2013) o, con más frecuencia, el sector institucional es una variable de estudio adicional, sin abordarse su estudio en profundidad, en trabajos con muy diversos objetivos (es el caso, por ejemplo, del artículo sobre las redes de colaboración en el sexto programa marco de Ortega y Aguillo, 2010).

En resumen, podemos señalar que los análisis comparativos de la actividad de los distintos sectores institucionales realizados a través de indicadores bibliométricos no son muy abundantes. En lo que se refiere a España, tiene especial interés el estudio antes mencionado de Moya-Anegón y otros (2013), que dedica un capítulo a los sectores institucionales en España, pero considera sólo cuatro sectores (frente a los seis estudiados en este trabajo), y analiza un periodo temporal más corto. El presente estudio analiza la producción científica de España por sectores institucionales para caracterizar su actividad y su evolución a lo largo de doce años. El interés de este tipo de agrupación es que los centros de un mismo sector institucional suelen compartir una misión, un contexto organizativo y unas políticas de gestión de recursos, lo que le confiere especial valor como unidad de análisis.

3. METODOLOGÍA

Para la realización de este estudio se ha descargado la producción científica española incluida en la base de datos WoS (SCIE, SSCI, A&HCI) en el periodo comprendido entre los años 2000 y 2011, utilizando el campo país (CU) y el término *Spain*, como criterios de búsqueda. La búsqueda se reali-

zó en 2014, siendo 2011 el año más reciente analizado para poder contar con una ventana fija de citación de tres años para todos los documentos. A través de un proceso semiautomático de tratamiento de datos se asignó a cada centro firmante incluido en los documentos su sector institucional correspondiente (Morillo y otros, 2013).

La clasificación de sectores institucionales utilizada distingue las siguientes seis agrupaciones:

- Administración. Incluye ministerios, gobiernos regionales, diputaciones, ayuntamientos y organismos autónomos vinculados a las diferentes administraciones, siempre y cuando no sean Organismos Públicos de Investigación.
- Empresas, tanto públicas como privadas.
- Entidades sin Ánimo de Lucro (ESAL). Comprende dos subgrupos: Entidades Tradicionales sin Ánimo de Lucro (ETSAL), que incluye asociaciones, colegios profesionales, fundaciones, sociedades científicas, parques, plataformas científico-tecnológicas, sindicatos, iglesia, etc.; y Estructuras de Investigación sin Ánimo de Lucro (EISAL), que agrupa redes de investigación, corporaciones, y entidades multisectoriales.
- Organismos Públicos de Investigación (OPI), que incluye el CSIC y otros organismos recogidos en la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación¹.
- Sector Sanitario. Comprende hospitales, centros de salud y centros socio-sanitarios. Debe tenerse en cuenta que en el caso de los hospitales universitarios, cuando firma un hospital o centro de salud junto con alguna universidad a la que están vinculados, se adjudica el documento al centro sanitario y no a la Universidad.
- Universidad.

Esta clasificación difiere de la propuesta por el Manual de Frascati, que distingue sólo cuatro agrupaciones (Administración Pública, Empresas, Enseñanza Superior e Instituciones Privadas sin Fines de Lucro), pero se ajusta mejor a los distintos tipos de centros que realizan investigación en España y es muy similar a la utilizada por otros autores (Maltrás y Quintanilla, 1992). Se ha considerado adecuada la separación de los Organismos Públicos de Investigación del resto de la Administración en tanto que su funcionamiento, financiación y finalidad son diferentes del resto de organismos que constituyen la jerarquía burocrática del Estado.

Hay que señalar que en un pequeño porcentaje de documentos, que no llega al 1% y que disminuye con los años, no se ha podido asignar un sector institucional, al no incluir la dirección de trabajo informacional suficiente, por lo que no se han considerado en el estudio. Aunque algunos datos iniciales se muestran para todo tipo documental, la mayoría de los resultados se presentan para "ítems citables", que incluye artículos originales y revisiones, agrupados bajo la denominación de "artículos".

Para los distintos sectores se incluyen distintos indicadores absolutos y relativos sobre su actividad, visibilidad y colaboración que se describen a continuación.

- *Indicadores de actividad*

Se cuantifica la actividad de los distintos sectores a través del número de publicaciones recogidas en WoS durante el período en estudio. Se han considerado 10 grandes áreas temáticas² –similares a las establecidas por el *Current Contents*– por agregación de las disciplinas en las que Thomson Reuters clasifica las revistas incluidas en WoS.

La especialización temática de los sectores se analiza mediante el índice de actividad (IA) que resulta del cociente entre el porcentaje de documentos que un sector dedica a un área y el porcentaje que dedica el total del país. Valores superiores a la unidad indican la mayor especialización de un sector en las áreas correspondientes.

Se ha utilizado la curva de Lorenz y el coeficiente de Gini para analizar la concentración de la producción por áreas temáticas de cada uno de los sectores. La curva de Lorenz permite describir la distribución relativa de una variable en un dominio determinado. Descrita en economía, permite medir las desigualdades de una sociedad al analizar las distribuciones de los ingresos y de la población, pero también se utiliza en otras disciplinas y con variables no propiamente económicas. El coeficiente de Gini es la expresión numérica del área que forma la curva de Lorenz con la bisectriz, que determina la igualdad perfecta en la distribución de la variable analizada. El coeficiente de Gini oscila entre 0 (distribución muy equitativa) y 1 (desigualdad total). Para su representación y cálculo se ha utilizado la hoja de cálculo de Barlan (2010).

- *Indicadores de impacto*

a) Los indicadores de impacto basados en revistas se calculan a partir de los datos del Factor de Impacto publicados en el *Journal Citation Reports* (JCR), que se considera indicador del prestigio y/o visibilidad de las revistas de publicación.

Se calculan los siguientes indicadores: porcentaje de artículos en revistas del primer cuartil (Q1) de cada disciplina en orden descendente de factor de impacto, y posición normalizada (PN) (Bordons y Barrigón, 1992). Debido a que Thomson no calcula factores de impacto para las revistas de Humanidades, este indicador no se usa en estas áreas.

b) Los indicadores de impacto a nivel documento estudian las citas obtenidas por los mismos, como una medida de su reconocimiento y utilidad para la creación de nuevo conocimiento. Si bien es cierto que los indicadores basados en citas han sido ampliamente utilizados y aceptados, no hay que olvidar sus limitaciones (véase, por ejemplo, Moed, 2005). En este estudio se han calculado las citas con dos ventanas de citación, una fija de tres años y otra variable desde la publicación del artículo hasta febrero de 2014. Sin embargo, sus resultados son similares por lo que sólo se muestran los obtenidos con la ventana fija. Se consideran los siguientes indicadores: número medio de citas por artículo, porcentaje de artículos no citados, porcentaje de artículos muy citados (HCP, siglas que corresponden a "*highly cited papers*") y citas relativas a España (CRE).

Para el cálculo del porcentaje de artículos muy citados se ha estudiado la contribución de los distintos sectores institucionales al 1% de los artículos de España más citados en cada disciplina. Hay que tener en cuenta que en algunas disciplinas varios artículos pueden tener el mismo número de citas que el umbral establecido para identificar el núcleo del 1% de artículos más citados; en estos casos todos ellos se consideran HCP, por lo que el porcentaje de artículos muy citados puede ser ligeramente superior al 1%. Por otro lado, si un artículo es considerado HCP en una disciplina, se considera como tal, independientemente de que pueda estar además asignado a otra u otras disciplinas donde no lo sea. Por último, señalar que para este cálculo se consideran sólo aquellas disciplinas que tienen 50 o más artículos en el período estudiado, y que se prescinde de los artículos de Humanidades, por la menor validez de las citas a corto plazo en estas disciplinas.

La tasa de citación relativa a España (CRE) compara las citas recibidas por un artículo con las citas medias de la producción del país en la disciplina a la que pertenece la revista de publicación del artículo y en el año de publicación. En el caso de que un artículo esté asignado a más de una disciplina, se calcula el indicador en cada una de ellas y posteriormente se obtiene la media. Finalmente, se calcula el valor medio para cada sector institucional, de forma que si el valor obtenido es superior a

uno, se puede concluir que su visibilidad está por encima de la del promedio del país.

- *Colaboración científica*

La colaboración científica se analiza a través de los siguientes indicadores: número de autores por artículo, número de centros por artículo y patrón de colaboración. Se denomina patrón de colaboración a la distribución de artículos según el tipo de colaboración entre instituciones: a) sin colaboración (una única institución); b) colaboración nacional: artículos en colaboración entre dos o más instituciones españolas; c) colaboración nacional e internacional: artículos con colaboración nacional e internacional (dos o más instituciones españolas y al menos una institución extranjera); y d) colaboración internacional: artículos en los que una institución española colabora con al menos una institución extranjera. Se analiza el patrón de colaboración de los diferentes sectores institucionales, su evolución a lo largo del tiempo, y la presencia de colaboración intersectorial.

Se utiliza el análisis de redes para representar la colaboración intersectorial. Los nodos representan a los sectores y su tamaño es proporcional a su producción. La colaboración entre sectores se representa mediante líneas, cuyo grosor será mayor cuanto más fuerte sea la relación. La fuerza de las relaciones viene dada por el índice Salton:

$$NDoc \text{ en colaboración entre sectores A y B}$$

$$\sqrt{NDoc \text{ en colaboración sector A} \times NDoc \text{ en colaboración sector B}}$$

La representación gráfica se realiza a través del software Pajek (Batagelj y Mrvar, 2014), aplicando el algoritmo Kamada-Kawai (Kamada y Kawai, 1989) y configurando el programa para que trate el peso de las relaciones como similitud entre nodos, mostrándose aquellos nodos con colaboraciones más fuertes más próximos entre sí.

4. RESULTADOS

La producción científica española ha aumentado desde 26.549 documentos en 2000 (2,2% de la producción mundial) hasta 61.230 en 2011 (3,4%), lo que supone un incremento de más de un 125%, aumento superior al observado para el total de la producción mundial en el mismo periodo (50%)³. Limitándonos a artículos originales y revisiones, que denominaremos "artículos" a lo largo del trabajo, la producción de España entre 2000 y 2011 asciende a 411.602 artículos, lo que representa el 82% del total de los 500.712 documentos del país.

El sector institucional con mayor producción en España es la Universidad (66%), seguida de los OPI (23%, correspondiendo el 20% al CSIC), el Sector Sanitario (18%), las Entidades sin Ánimo de Lucro (ESAL) (10%, de las cuales el 6% corresponde a las EISAL y el 4% restante a las ETSAL), la Administración⁴ (4%) y las Empresas (4%). La contribución de los distintos sectores a lo largo de los años se mantiene bastante estable, a excepción de las ESAL, que experimentan un importante incremento de forma que su aportación aumenta desde el 3% en 2000 al 17% en 2011 (la producción de las EISAL pasa del 2% al 12% y la de las ETSAL del 1% al 7%).

Mención especial merecen las EISAL, que incluyen un grupo heterogéneo de entidades, de carácter multisectorial, que han incrementado de forma sustancial su producción a lo largo del periodo. Incluyen algunos institutos de investigación médica, como el Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), formado por numerosos grupos de investigación de distintas entidades catalanas; entidades de gestión de la investigación, como la Institució Catalana de Recerca i Estudis Avançats (ICREA), que contrata investigadores para su integración en equipos de investigación de universidades y centros de investigación; o los Centros de Investigación Biomédica en Red (CIBER), que promueven la colaboración entre grupos consorciados y aportan en el periodo estudiado el 23% de la producción de las ESAL.

- *Especialización temática de los sectores institucionales*

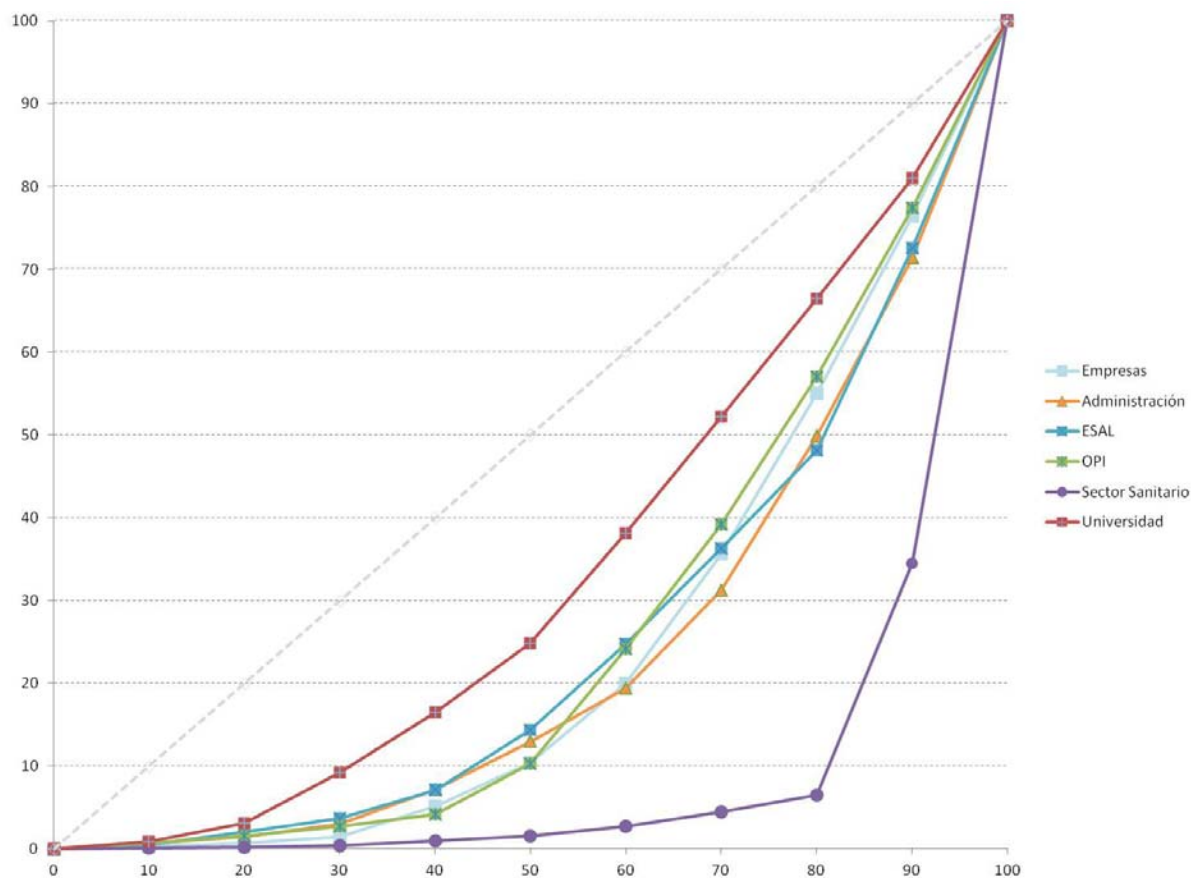
La distribución porcentual de la producción de cada sector institucional por áreas temáticas ofrece una clara imagen de su perfil de actividad. El perfil de la Universidad es muy parecido al del total del país por el importante peso de la producción de este sector en el total de la producción nacional (tabla I).

La Universidad es, además, el sector que muestra una actividad más distribuida por áreas temáticas, tal y como demuestra su menor coeficiente de Gini (0,32). En el otro extremo se sitúa el Sector Sanitario, que concentra el 93% de su producción en solo dos áreas: Medicina Clínica y Biomedicina (coeficiente de Gini=0,80). La figura 1 muestra la distribución relativa de artículos por disciplinas para los distintos sectores institucionales (curva de Lorenz), que es la expresión gráfica del coeficiente de Gini. La bisectriz de la figura corresponde a la situación de perfecta equidad (coeficiente de Gini=0), y la de mayor concentración corresponde al Sector Sanitario que muestra la curva más alejada de la bisectriz de la figura y, por consiguiente, una mayor área entre la bisectriz y la curva.

Tabla I. Perfil de actividad temática de los sectores institucionales españoles y del total del país (distribución porcentual de la producción por áreas) (WoS 2000–2011)

Sector institucional	Agric., Biol., Medio Ambiente	Biomedicina	Ciencias Sociales	Física	Humanidades	Ingeniería, Tecnología	Matemáticas	Medicina Clínica	Multidisciplinar	Química
Administración	35,42	14,58	7,12	23,04	1,84	8,01	0,86	26,62	0,96	5,20
Empresas	31,16	28,33	4,79	6,96	0,30	20,78	1,03	25,68	0,63	12,66
ESAL	13,45	31,76	4,46	14,91	0,61	15,36	1,96	35,53	2,18	9,42
EISAL	10,22	32,84	3,95	17,80	0,42	11,75	1,78	39,11	2,47	8,45
ETSAL	17,68	33,11	5,13	9,55	0,82	19,67	2,03	30,15	2,14	11,19
OPI	26,76	23,43	1,90	29,62	1,05	18,20	0,92	8,02	1,56	19,80
Sector Sanitario	2,18	34,33	2,53	0,20	0,09	1,35	0,14	80,36	0,72	0,78
Universidad	18,32	17,28	9,43	18,96	2,86	24,77	7,98	10,89	1,16	18,59
Total España	17,85	20,56	7,02	17,64	2,22	20,10	5,49	21,14	1,17	15,73

Figura 1. Curva de Lorenz o grado de concentración de la producción científica de los sectores institucionales españoles por áreas (WoS 2000–2011)



La especialización temática de los diferentes sectores institucionales se presenta en la tabla II, mediante el cociente de la proporción de artículos que cada sector dedica a un área frente al porcentaje que dedica el total de España. El Sector Sanitario es el que muestra la mayor especialización, concretamente en Medicina Clínica, aunque también otros sectores presentan cierta especialización en esta área – ESAL (especialmente las EISAL), Empresas y Administración Pública–, frente a la escasa dedicación al tema de otros como los OPI y la Universidad. Destaca la alta especialización de la Administración, Empresas y OPI en Agricultura/Biología/Medio Ambiente. La Administración muestra también una alta actividad relativa en Física⁵, los OPI en Física y Química, y las Empresas en Biomedicina y Medicina Clínica, áreas en las que también destacan las ESAL. Finalmente, señalar que la Universidad presenta una mayor dedicación que el promedio del país en Matemáticas, Ciencias Sociales, Humanidades, Ingeniería/Tecnología y Química.

A lo largo del periodo la mayor parte de los sectores mantienen su especialización temática. Los cambios más destacables entre el primer y el último bienio corresponden al sector ESAL, que muestra una creciente especialización en el área Multidisciplinar (+0,57 puntos entre los dos bienios analizados), Química (+0,54) y Biomedicina (+0,41), y una menor especialización en Medicina Clínica (-0,54) y Ciencias Sociales (-0,53). La Administración pierde especialización en Física (-0,37), aunque mantiene un índice relativo

a España en el último periodo considerablemente alto (1,32). También es destacable que el Sector Sanitario aumenta su especialización en revistas del área Multidisciplinar (+0,39) y en Biomedicina (+0,32).

- *Impacto de los sectores institucionales: indicadores absolutos y relativos*

La tabla III muestra un resumen de la actividad científica y el impacto de los distintos sectores institucionales, así como su comparación con la media nacional. Los OPI y las ESAL, que constituyen el 23% y el 10% de la producción española respectivamente, son los dos sectores que obtienen los valores más altos en todos los indicadores relativos de impacto. Destacan especialmente las EISAL con un 44% más de artículos en primer cuartil y un valor de citas por artículo un 35% más elevado que el promedio del país.

La comparación entre el primer y el último bienio de los indicadores de actividad e impacto relativo de cada sector (tabla IV) permite observar el ya mencionado incremento del porcentaje de los artículos publicados por las ESAL, con un aumento de 13 puntos porcentuales – 10 puntos en el caso de las EISAL y 5 en el caso de las ETSAL – y la reducción en el porcentaje de artículos no citados con respecto a España tanto de las ESAL como del Sector Sanitario y los OPI. Los OPI y las ESAL, especialmente las EISAL, presentan los valores más altos de citas relativas en ambos bienios. A lo largo de los años se observan cambios muy ligeros en las citas relativas de los distintos sectores.

Tabla II. Especialización temática de los sectores institucionales españoles a través de su índice de actividad (WoS 2000–2011)

Sector institucional	Agric., Biol., Medio Ambiente	Biomedicina	Ciencias Sociales	Física	Humanidades	Ingeniería, Tecnología	Matemáticas	Medicina Clínica	Multidisciplinar	Química
Administración	1,98	0,71	1,01	1,31	0,83	0,40	0,16	1,26	0,82	0,33
Empresas	1,75	1,38	0,68	0,39	--	1,03	0,19	1,21	0,53	0,80
ESAL	0,75	1,54	0,63	0,85	0,28	0,76	0,36	1,68	1,86	0,60
EISAL	0,57	1,60	0,56	1,01	0,19	0,58	0,32	1,85	2,11	0,54
ETSAL	0,99	1,61	0,73	0,54	0,37	0,98	0,37	1,43	1,83	0,71
OPI	1,50	1,14	0,27	1,68	0,47	0,91	0,17	0,38	1,34	1,26
Sector Sanitario	0,12	1,67	0,36	0,01	--	0,07	0,02	3,80	0,61	0,05
Universidad	1,03	0,84	1,34	1,07	1,29	1,23	1,45	0,52	0,99	1,18

Nota: se muestra el índice de actividad de los sectores en aquellas áreas con más de 100 artículos en el periodo analizado.

Tabla III. Actividad e impacto de los sectores institucionales españoles: indicadores absolutos y relativos al total del país (WoS 2000–2011)

Sector institucional	Datos de sectores					Indicadores relativos a España				
	Art	Citas/Art	% Art sin citas	PN	% Art Q1	% Art	Citas relativas	Indicador relativo de no citación	Relativo PN	Relativo % Art Q1
Administración	17245	5,55	20,64	0,66	45,42	4,19	0,93	0,95	1,00	0,98
Empresas	15653	4,47	22,12	0,63	41,80	3,80	0,87	1,02	0,95	0,91
ESAL	39312	9,23	10,96	0,74	59,48	9,55	1,29	0,51	1,12	1,29
EISAL	25037	10,48	8,39	0,76	66,28	6,08	1,35	0,39	1,15	1,44
ETSAL	16996	7,83	13,77	0,71	56,26	4,13	1,10	0,63	1,08	1,22
OPI	92770	7,46	12,73	0,74	60,80	22,54	1,13	0,59	1,12	1,32
Sector Sanitario	74236	6,63	21,29	0,59	37,29	18,04	0,95	0,98	0,89	0,81
Universidad	271382	4,71	23,39	0,66	44,97	65,93	0,94	1,08	1,00	0,98

Tabla IV. Evolución del porcentaje de artículos de los sectores institucionales españoles e indicadores de impacto relativos a España (WoS 2000–2001 y 2010–2011)

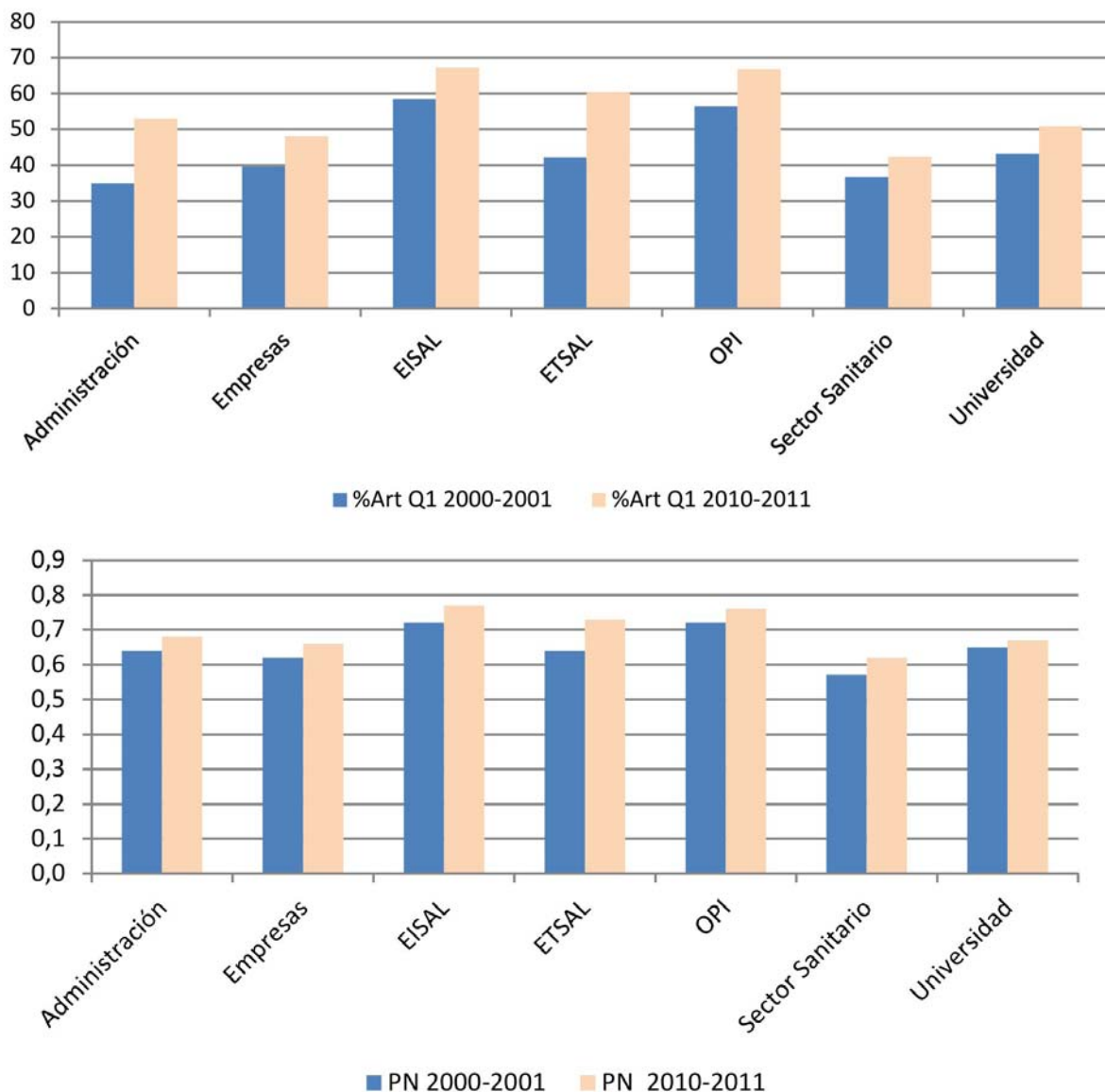
Sector institucional	2000-2001			2010-2011			Diferencias		
	%	Citas Relativas	Índice relativo no citación	%	Citas Relativas	Índice relativo no citación	%	Citas Relativas	Índice relativo no citación
Administración	3,16	0,92	0,96	5,18	0,89	0,99	2,02	-0,03	0,03
Empresas	3,17	0,91	1,04	4,09	0,85	1,04	0,93	-0,06	0,00
ESAL	3,17	1,23	0,76	16,45	1,26	0,48	13,28	0,03	-0,28
EISAL	1,72	1,31	0,49	11,29	1,35	0,38	9,57	0,04	-0,12
ETSAL	1,47	1,13	1,07	6,78	1,12	0,59	5,31	-0,01	-0,48
OPI	22,53	1,18	0,68	22,80	1,13	0,49	0,28	-0,05	-0,20
Sector Sanitario	19,12	0,97	1,09	17,65	0,96	0,93	-1,47	-0,01	-0,17
Universidad	64,36	0,96	1,02	66,52	0,91	1,11	2,15	-0,05	0,09

A lo largo de los años aumenta el porcentaje de artículos que publica España en las mejores revistas de cada disciplina (Q1), que asciende desde el 43% en 2000 hasta el 51% en 2011, destacando los valores superiores a la media del país a lo largo de todo el periodo de OPI y ESAL. La tendencia ascendente en el porcentaje de artículos Q1 se verifica en los distintos sectores institucionales, observándose incrementos entre el primer y el último bienio que van desde 6 puntos porcentuales en el Sector Sanitario hasta 18 puntos en la Administración y las ETSAL. A lo largo del periodo se observan también incrementos en la Posición Normaliza-

da (PN) de las revistas de publicación en todos los sectores, aunque en este caso las diferencias son muy ligeras. En cualquier caso, se verifica a través de los indicadores una tendencia creciente a publicar en revistas de mayor impacto (figura 2).

Considerando el 1% de los artículos españoles más citados en las distintas disciplinas WoS, se ha analizado en qué proporción han participado cada una de las agrupaciones institucionales (tabla V). Resulta muy interesante observar que los Organismos Públicos de Investigación y las Entidades sin Ánimo de Lucro, así como sus dos subsecto-

Figura 2. Evolución del porcentaje de artículos en revistas del primer cuartil (Q1) y de la Posición Normalizada (PN) de los sectores institucionales españoles (WoS 2000–2001, 2010–2011)



res, tienen una mayor proporción de HCP de lo que sería esperable en función de su número de artículos. En el total del periodo, las ESAL producen casi el 10% de los artículos españoles, pero son responsables del 19% de los HCP (aunque las EISAL alcanzan mayores valores relativos de HCP que las ETSAL), mientras que los OPI producen el 23% de los artículos del país y el 28% de los HCP. Es decir, que la contribución de las ESAL y de los OPI a los HCP de España es un 96% y un 28% –respectivamente– más elevada de lo esperado. La Universidad, sin embargo, y sobre todo, la Administración y las Empresas, están infrarrepresentadas en el conjunto de HCP atendiendo a

su producción total, lo que sugiere una menor repercusión de sus publicaciones.

A lo largo de los años, la contribución de los distintos sectores al núcleo del 1% de publicaciones de España más citadas en cada disciplina varía, aumentando la contribución de la Administración y las ESAL (en especial las EISAL) mientras que se reduce la de los restantes sectores. Como se ha comentado antes, la contribución de los OPI y las ESAL a los HCP es superior a lo que sería esperable en función de su producción, pero mientras que las ESAL tienden a incrementar su contribución a los HCP a lo largo del tiempo, los OPI reducen su participación.

Tabla V. Distribución comparativa del porcentaje de HCP y artículos de España por sectores institucionales (WoS 2000–2011 y bienios 2000–2001 y 2010–2011)

Sector institucional	2000–2001			2010–2011			2000–2011		
	%HCP	%Art	%HCP/ %Art	%HCP	%Art	%HCP/ %Art	%HCP	%Art	%HCP/ %Art
Administración	2,38	3,16	0,75	3,38	5,18	0,65	3,08	4,19	0,74
Empresas	3,06	3,17	0,97	3,05	4,09	0,74	2,60	3,80	0,68
ESAL	4,24	3,17	1,34	27,50	16,45	1,67	18,67	9,55	1,96
EISAL	1,87	1,72	1,08	21,07	11,29	1,87	13,28	6,08	2,18
ETSAL	2,55	1,47	1,73	10,15	6,78	1,50	7,23	4,13	1,75
OPI	31,92	22,53	1,42	30,12	22,80	1,32	28,81	22,54	1,28
Sector Sanitario	20,20	19,12	1,06	18,19	17,65	1,03	17,81	18,04	0,99
Universidad	55,52	64,36	0,86	52,71	66,52	0,79	55,66	65,93	0,84

Impacto por áreas temáticas

¿Obtienen los distintos sectores un mayor impacto en aquellas áreas científicas en las que están más especializados? Para responder a esta cuestión, se comparan las citas obtenidas por los artículos de cada uno de los sectores, respecto a los valores medios de citas del conjunto nacional en las disciplinas de publicación de los artículos (tabla VI), y se estudia si los valores más altos corresponden a sus campos de especialización (descritos en la tabla II).

Nuestros resultados muestran que los sectores no obtienen su mayor impacto relativo en las áreas en las que están especializados. El hallazgo más llamativo es que los OPI y, sobre todo, las ESAL obtienen los valores más altos de citas relativas en todas las áreas, siendo sus tasas de citación superiores a la media del país en todos los casos – sólo en el caso de la Ingeniería/Tecnología las ETSAL presentan valores sensiblemente inferiores al promedio nacional –. Así, aunque las ESAL están especializadas en Biomedicina (32%), no es aquí donde obtienen sus mayores índices relativos de citación, sino en otras áreas como Física y Ciencias Sociales, en las que tienen escasa producción. También son reseñables los valores de citas relativas obtenidos por los Organismos Públicos de Investigación en Ciencias Sociales –cuya actividad en el área no alcanza el 2% de su producción total–. Por otra parte, sectores como las Empresas, con una dedicación importante en Biomedicina (29%) o Ingeniería/Tecnología (17%), presentan en estas áreas unos índices de impacto en torno a un 20% por debajo de la media nacional. Resulta notorio el alto impacto relativo del Sector Sanitario en el área Multidisciplinar, un 70% por encima de la media del

país, pero hay que recordar que corresponde a un bajo porcentaje de artículos (menos del 1%).

- *Colaboración de los sectores institucionales*

El tamaño del equipo de investigación de los distintos sectores institucionales varía entre los 7 autores de media de los artículos publicados por las Empresas y los 17 de los Organismos Públicos de Investigación, mientras que el número medio de centros oscila entre los 3 y 5 organismos (figura 3)⁶. Los valores medios de España corresponden a 8 autores por artículo y a 3 centros por artículo. Sin embargo, las diferencias entre sectores se reducen si se consideran las medianas (entre 4 y 6 autores y entre 2 y 4 centros), que eliminan el efecto de los valores atípicos (algunos documentos con un número de autores o centros muy elevado). Tanto el número medio de autores como el de centros experimenta un importante crecimiento entre el primer y el último bienio, fundamentalmente en las ESAL y los OPI, cuyo número medio de organismos por artículo se duplica y el número medio de autores se triplica.

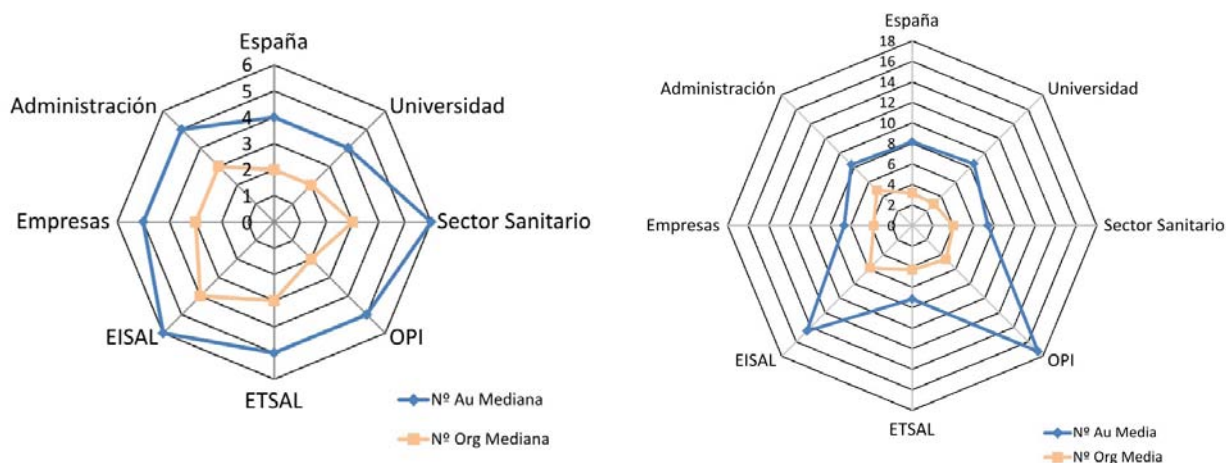
Un tercio de los artículos de España están firmados por un único centro, estando la colaboración nacional presente en el 40% de los artículos y la internacional en el 37% del total (figura 4). Este patrón de colaboración es muy similar al que presenta la Universidad, principal sector institucional productor de artículos, pero se observan diferencias entre sectores. Así, el porcentaje de los artículos realizados por un solo centro oscila entre el 8% de las ESAL y el 32% de la Universidad. En lo que respecta a la colaboración internacional (colaboración sólo internacional más colaboración nacional e internacional), los OPI y las ESAL alcanzan los valores más altos (45%

Tabla VI. Citas relativas a España de los sectores institucionales españoles por áreas temáticas (WoS 2000–2011)

Sector institucional	Agric., Biol., Medio Ambiente	Biomedicina	Ciencias Sociales	Física	Ingeniería, Tecnología	Matemáticas	Medicina Clínica	Multidisciplinar	Química
Administración	0,90	0,96	0,97	1,04	0,79	0,58	0,86	1,32	0,88
Empresas	0,95	0,85	1,09	0,85	0,75	0,75	0,93	0,85	0,86
ESAL	1,27	1,26	1,46	1,42	1,14	1,26	1,31	1,47	1,27
EISAL	1,38	1,28	1,71	1,57	1,40	1,27	1,36	1,53	1,42
ETSAL	1,20	1,27	1,18	1,04	0,95	1,24	1,20	1,48	1,10
OPI	1,12	1,20	1,29	1,15	1,10	1,21	1,17	1,28	1,10
Sector Sanitario	0,99	1,00	1,03	1,03	0,87	0,65	0,94	1,80	0,83
Universidad	0,89	0,99	0,89	0,96	0,94	0,93	0,94	0,77	1,02

Nota: Se destacan los valores de citas relativas superiores a 1,10. No se muestran las Humanidades por la menor validez de las citas a corto plazo en esta área.

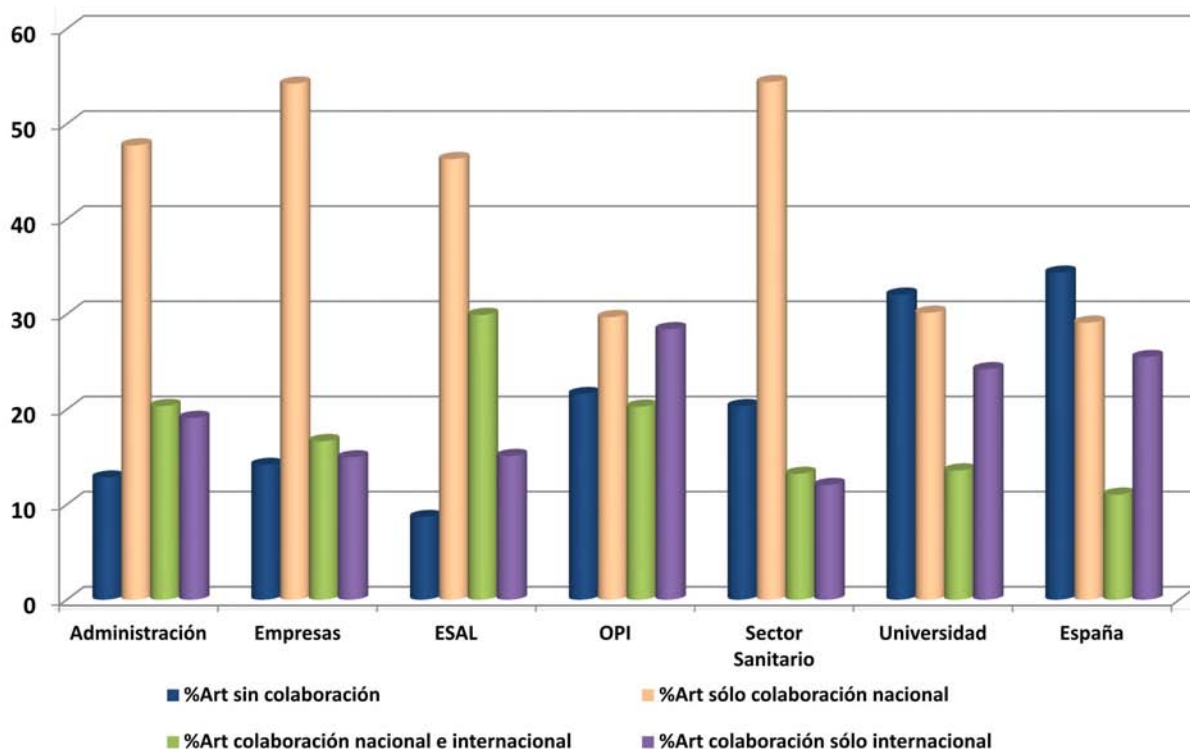
Figura 3. Mediana y media del número de autores y organismos de los sectores institucionales españoles (WoS 2000–2011)



y 49% respectivamente), mientras que el Sector Sanitario y las Empresas muestran valores inferiores a la media del país (25% y 32% respectivamente). La colaboración nacional (suma de colaboración sólo nacional y colaboración nacional e internacional) es superior a la del conjunto de la producción española en todos los sectores analizados excepto en la Universidad, destacando los valores obtenidos por las ESAL (76%) y las Empresas (71%).

La comparación del bienio inicial y final del periodo pone de manifiesto los principales cambios que se producen en las prácticas de colaboración a lo largo de los años. En el total del país, se observa que el porcentaje de artículos realizados por un solo centro se reduce en 14 puntos porcentuales, mientras que se incrementa la fracción de producción en colaboración nacional (+11 puntos) e internacional (+12 puntos). El descenso relati-

Figura 4. Patrón de colaboración de los sectores institucionales españoles (WoS 2000–2011)



vo de la producción realizada por un solo centro se evidencia en todos los sectores institucionales, aunque es menor en la Administración y máximo en el Sector Sanitario. El incremento de la colaboración nacional es similar en los distintos sectores, mientras que el mayor incremento de la colaboración internacional se observa en el sector de las ESAL (+20 puntos).

El estudio de las relaciones intersectoriales permite apreciar cómo los diferentes sectores colaboran fundamentalmente con la Universidad (tabla VII), si bien son especialmente destacables los valores obtenidos por las Empresas, las ESAL y la Administración. Por el contrario, la menor colaboración se produce siempre con las Empresas, de forma que ningún sector publica más del 8% de sus artículos con la industria. En cuanto a la colaboración entre instituciones del mismo sector, es el Sanitario el que alcanza un mayor porcentaje (35% de los artículos en colaboración), junto con la Universidad (33%), mientras que la menor colaboración intra-sectorial corresponde a las Empresas (7%).

La colaboración Universidad–Empresa, que es objeto de numerosos estudios en la literatura por

su importante papel en la transferencia de conocimiento entre el sector público y el privado, supone el 4% de los documentos en colaboración de la Universidad y el 51% de los correspondientes a las Empresas, y ha experimentado un ligero aumento a lo largo del periodo. En el bienio inicial, los artículos en colaboración entre estos sectores representaban el 1% del total de los artículos españoles, alcanzando el 2% en el bienio final.

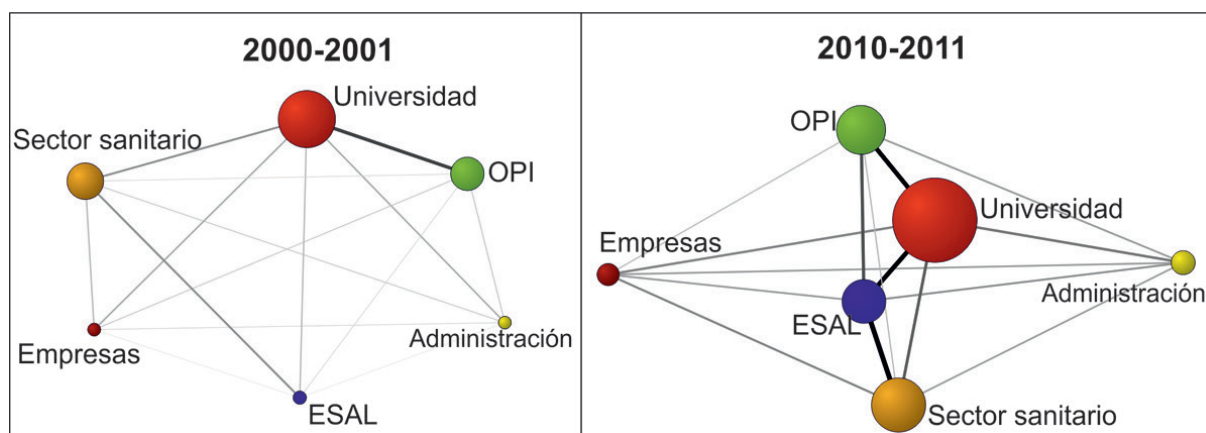
La figura 5 muestra cómo la producción en colaboración entre los distintos sectores (nodos) y la fuerza de los vínculos entre ellos (grosor de las líneas) aumenta a lo largo del tiempo. Destaca la evolución de las ESAL, que en sus inicios mostraban una baja producción en colaboración ligada en su mayor parte al Sector Sanitario, y en los últimos años, no sólo ha aumentado considerablemente su producción en colaboración, sino que además se aprecia una mayor intensidad en sus relaciones, sobre todo con la Universidad y los OPI, manteniendo la fuerte colaboración inicial con el Sector Sanitario. La Universidad y las ESAL desempeñan un importante papel en el núcleo de la red en el segundo periodo, en el que se crean dos importantes triadas que incluyen a los OPI y al Sector Sanitario.

Tabla VII. Porcentaje de colaboración intersectorial de la producción científica española (WoS 2000–2011)

Sector Institucional	Administración	Empresas	EISAL	ETSAL	OPI	Sector Sanitario	Universidad
Administración	10,56	8,49	4,63	6,19	4,31	4,64	3,85
Empresas	7,59	7,23	4,23	5,46	2,85	5,15	3,69
EISAL	7,21	7,38	12,19	17,91	6,67	12,13	5,85
ETSAL	6,25	6,17	11,61	4,44	3,09	6,29	4,15
OPI	20,85	15,45	20,71	14,81	15,42	7,49	16,52
Sector Sanitario	18,25	22,65	30,62	24,48	6,09	35,25	9,93
Universidad	47,26	50,70	46,04	50,43	41,87	30,97	33,20
Total art. colab.	15.034	13.437	23.430	15.190	72.736	59.144	184.399

Nota: Porcentajes calculados en columnas sobre el total de artículos en colaboración de cada sector institucional. La diagonal incluye la colaboración intra-sectorial.

Figura 5. Colaboración entre los sectores institucionales españoles (WoS 2000–2001 y 2010–2011)



5. DISCUSIÓN Y CONCLUSIONES

La producción científica española en los primeros doce años del siglo XXI sigue la tendencia ascendente iniciada en la mitad de la década de los ochenta en respuesta a un intento de estructuración e impulso de la investigación nacional (Bordons, 2004; Maltrás y Quintanilla, 1992). De esta manera, en este estudio se observa: un importante incremento en el número de publicaciones, en especial de algunos sectores como las ESAL; una tendencia creciente a publicar en revistas de alto impacto; un desarrollo progresivo de prácticas colaborativas en la investigación; y una intensificación de su internacionalización a través de una mayor colaboración con investigadores de centros extranjeros. No obstante, la participación del sector Empresas en la investigación continua siendo muy reducida.

El sector universitario es responsable de dos terceras partes de la producción científica a lo lar-

go de todo el periodo. Este importante papel de la Universidad en la generación de nuevo conocimiento y en su difusión a través de publicaciones científicas coincide con lo descrito en otros países, como Holanda (80% de la producción procede de la Universidad), Estados Unidos (75%) o el Reino Unido (60%) (Tijssen y otros, 2010; National Science Board, 2014; Katz y Hicks, 1996). En general, la aportación relativa de la Universidad es algo menor en aquellos países que cuentan con un sector fuerte de centros públicos de investigación (OPI), como Italia –donde cuentan con el *Consiglio Nazionale delle Ricerche* (CNR)–, Alemania –la *Max Planck Society* o la *Fraunhofer Society*–, o Francia –el *Centre National de la Recherche Scientifique* (CNRS)–. En España el mayor OPI es el Consejo Superior de Investigaciones Científicas (CSIC), que aporta el 20% de la producción de España y el 88% de la del sector OPI. Su peso es inferior al correspondiente al CNRS (36%), pero superior al del CNR (12%) o la Max Plack (10%) en sus respectivos

países (Grupo de Bibliometría, 2006). En lo que se refiere al sector hospitalario, éste aporta el 18% de la producción de España, porcentaje superior al descrito en otros países como los Países Bajos (8%, Tijssen y otros, 2010), pero hay que tener en cuenta que en nuestro caso se contabiliza la producción de los hospitales universitarios, que en el informe holandés se asigna al sector universitario. El sector empresarial muestra poca actividad en España (4%), por debajo de la registrada en otros países como los Estados Unidos (6%, National Science Board, 2014) o los Países Bajos (8%, Tijssen y otros, 2010).

En lo que respecta al perfil temático de actividad de los distintos sectores, resulta lógica la alta concentración de la producción del Sector Sanitario en Medicina Clínica, área a la que este sector dedica casi cuatro veces más documentos que el promedio del país. En el otro extremo se sitúa la Universidad, que es el sector que presenta mayor dispersión temática en su perfil de actividad y el único que muestra alta especialización en algunas áreas como Ciencias Sociales, Humanidades y Matemáticas; lo que pone de manifiesto la importancia de esta institución para el avance del conocimiento en los mencionados campos. Por otra parte, la alta dedicación de las ESAL a la Medicina Clínica y la Biomedicina se debe a que estas áreas han sido especialmente sensibles a la creación de redes y consorcios, pudiéndose destacar distintas entidades de alta producción como IDIBAPS o CIBER. La alta especialización de la Administración en Agricultura/Biología/Medio Ambiente se explica por la actividad de distintos centros dependientes de la Administración central y, sobre todo, autonómica. Es el caso del Instituto Valenciano de Investigaciones Agrarias (IVIA) o el Instituto Andaluz de Investigación Agraria y Pesquera (IFAPA), que realizan investigación de especial interés para sus regiones. Por lo que se refiere a los OPI, su especialización viene marcada por la orientación temática de los centros del CSIC, que presenta más actividad que el promedio del país en Física, Agricultura/Biología/Medio Ambiente y Química (Bordons y otros, 2014).

Auge de las Estructuras de Investigación sin Ánimo de Lucro (EISAL)

En el periodo en estudio, los distintos sectores institucionales aumentan su producción de forma proporcional a la del total nacional, con la única excepción de las Entidades sin Ánimo de Lucro, que experimentan un crecimiento muy superior al promedio nacional, en especial las EISAL, que suponen el 64% de los artículos de las ESAL en todo el periodo (55% en el primer bienio vs 69% en el último).

En este estudio hemos denominado EISAL a un grupo heterogéneo de entidades, muchas de ellas multisectoriales, que han surgido en muchos países impulsadas por los gobiernos para promover una investigación eficaz y de calidad entre grupos que pueden incluso estar físicamente alejados –en el caso de las redes– pero que trabajan en los mismos temas y pueden beneficiarse de estrategias específicas de financiación y gestión de recursos (Cruz-Castro y otros, 2012). Se ha detectado su aparición tanto en las economías emergentes (Adams, 2012) como en las consolidadas, y entre estas últimas se pueden mencionar los programas *Research Coordination Networks* de la *National Science Foundation* (http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=11691), las *Collaborative Research Networks* del *Department of Education* del Gobierno de Australia (<https://www.education.gov.au/collaborative-research-networks-crn>), o la *Canadian Network of Centres of Excellence* (<http://www.nce-rce.gc.ca/>). A través de estas iniciativas los gobiernos pretenden fomentar la interdisciplinariedad, poner énfasis no sólo en la excelencia sino también en el uso y la transferencia; y promover la cooperación entre el sector público y privado.

En el caso de España, la política científica de los últimos años ha intentado fortalecer la investigación cooperativa, en la que los diferentes agentes de la I+D+i, tanto en el ámbito nacional como en el internacional, persiguen una optimización de los recursos disponibles. En concreto, el Plan Nacional I+D+i 2008–2011 establecía la necesidad de impulsar “los programas destinados a actuaciones concertadas o cooperativas, como la creación de consorcios de I+D+i”, poniendo especial énfasis en la cooperación entre la investigación pública y privada, y en garantizar y promover la transferencia de conocimiento y de tecnología (Comisión Interministerial de Ciencia y Tecnología, 2007). Por otro lado, hay que mencionar los proyectos CIBER y RETICS, dentro del Programa Ingenio 2010, que han impulsado la investigación en estructuras duraderas de investigación cooperativa, contribuyendo a crear redes estables y multidisciplinares en el área de ciencias de la salud.

El sector Empresas

El sector empresarial es el que muestra menor actividad científica a través de los indicadores basados en publicaciones, aunque experimenta un crecimiento positivo en el periodo, ligeramente por encima del descrito para el total del país. Su actividad se centra en Agricultura/Biología/Medio Ambiente (31%); Biomedicina (28%) y Medicina Clínica (26%), siendo las industrias de cabecera

algunas empresas públicas del ámbito agroalimentario y sanitario –como el *Institut de Recerca i Tecnologia Agroalimentàries* o la Escuela Andaluza de Salud Pública– y empresas farmacéuticas –como Pfizer o Pharmamar–. El sector que más colabora con las Empresas es la Administración (7% de sus artículos en colaboración), seguido del Sector Sanitario y las ESAL (5% cada uno), mientras que la Universidad sólo colabora en el 4% de sus publicaciones. Este nivel de colaboración Universidad–Empresa puede considerarse medio–bajo atendiendo al estudio del Centre for Science and Technology Studies (CWTS) donde se describen las tasas de colaboración entre la Universidad y las Empresas para más de 700 universidades de todo el mundo (<http://www.cwts.nl/UIRC2014>), y en el que sólo 4 de las 28 universidades españolas recogidas muestran tasas por encima de este nivel. Ello a pesar de las iniciativas de los gobiernos españoles de impulsar la participación empresarial en el sistema de I+D+i.

Impacto de la investigación

Las ESAL y los OPI son los sectores institucionales que presentan valores más altos de impacto, tanto en términos de citas como de porcentaje de artículos muy citados. Esto es un signo de la relevancia de su investigación y puede verse favorecido por la orientación internacional de la misma y sus altas tasas de colaboración con investigadores extranjeros, ya que esta última se ha asociado a una mayor citación de las publicaciones (Glänzel y Schubert, 2001; Díaz–Faes y otros, 2013). A lo largo del periodo todos los sectores institucionales tienden a publicar un mayor porcentaje de su producción en revistas de alto impacto, pero son las ESAL, en particular las EISAL, las que muestran el mayor incremento en su capacidad de producir artículos muy citados. Este dato apoya el éxito de estas nuevas formas organizativas para producir investigación de excelencia.

Aunque el impacto del Sector Sanitario está por debajo de la media nacional, hay que tener en cuenta algunos factores importantes. Por un lado, el hecho de que algunos hospitales participan en entidades multisectoriales, y que, por tanto, parte de su investigación más puntera queda adscrita en este estudio a las ESAL. Por otro lado, hay que mencionar que este sector hace un importante uso de revistas españolas (21% de los artículos del sector sanitario, muy por encima del 9% del total del país)⁷. Muchas de estas revistas, escritas en español, están dirigidas sobre todo a la comunidad nacional y no suelen ocupar los primeros puestos en los rankings por Factor de Impacto del JCR (Van Leeuwen, 2001), pero tienen una importante fun-

ción apoyando la práctica médica y fomentando la investigación en los hospitales.

Contrariamente a lo que podría esperarse, no se observa correspondencia entre las áreas en las que los sectores institucionales obtienen más impacto y aquéllas en las que están más especializados. Así, las ESAL y los OPI, que muestran los valores más altos de impacto relativo, lo obtienen en casi todas las áreas temáticas. En el otro extremo se sitúa la Universidad, que sólo obtiene una tasa de citación por encima del promedio español en el área de Química. Estos datos hay que interpretarlos con cautela, ya que interviene el efecto “tamaño”, que hace que sea más difícil obtener altos indicadores relativos de impacto para unidades con un importante volumen de producción, porque los posibles “outliers” de alta citación quedan compensados con una gran masa de publicaciones escasamente citadas. Por esta razón, es más adecuado abordar el estudio de la relación entre especialización y citación a un menor nivel de agregación: por ejemplo, a nivel de centros o instituciones concretas.

Investigación colaborativa

El tamaño de los grupos de investigación, medido a través del número medio de autores firmantes de un artículo, viene en parte determinado por la especialización temática de los sectores (Wuchty y otros, 2007). Así, destaca el alto número medio de autores por artículo de los OPI, debido al importante peso de sus publicaciones en el área de Física (28% de la producción de los OPI), en el que la multiautoría es una constante debido a la utilización de grandes instalaciones internacionales (*Big Science*). En lo que se refiere al número de centros por artículo, las diferencias entre sectores son menos acusadas. Resulta llamativo que la Universidad presente el menor número medio de centros por artículo, que, en parte, puede atribuirse al peso de las Humanidades y Ciencias Sociales en este sector, aunque no puede descartarse la influencia de otros factores, ya que los artículos de estas áreas solo constituyen cerca del 12% del total de su producción. En este sentido, un mayor tamaño de los grupos o departamentos universitarios, que necesitarían recurrir a colaboraciones externas con menos frecuencia, podría ser una razón.

En lo que se refiere a las tasas de colaboración, y en línea con lo anterior, la Universidad presenta el mayor porcentaje de artículos firmados por un único centro. Predomina la colaboración nacional sobre la internacional en todos los sectores y, aunque tradicionalmente la colaboración nacional alcanzaba los valores más altos en el Sector Sanitario, por la colaboración intra- e inter-hospitalaria, ambas

frecuentes en el desarrollo de estudios o guías clínicas; otros sectores, como las ESAL, alcanzan cifras muy elevadas de este tipo de colaboración, probablemente por la presencia de entidades consorciadas como las mencionadas anteriormente.

En lo que se refiere a la colaboración internacional, destacan las ESAL y los OPI, en parte porque tienen actividad en áreas en las que este tipo de relaciones suelen ser habituales (Física), pero también alcanzan altas tasas de colaboración internacional en otras áreas, lo que pone de manifiesto su capacidad para atraer socios extranjeros y sugiere una mayor internacionalidad de sus temas de investigación. Los OPI y las ESAL presentan además el mayor número de centros por artículo, por lo que sus redes de colaboración probablemente son también más heterogéneas, lo que se ha revelado beneficioso para la investigación por la mayor creatividad asociada a la diversidad de enfoques y aproximaciones (Hackett, 2005).

A lo largo del periodo tiende a aumentar en todos los sectores, pero sobre todo en las ESAL y los OPI, la actividad en colaboración entre autores y centros, tanto dentro del propio país como con centros extranjeros. Los recortes en los presupuestos dedicados a la investigación, que favorecen la colaboración para compartir recursos; la creciente importancia de los campos interdisciplinares que necesitan la integración de personal de diferentes ámbitos científicos; y los intereses culturales, económicos o geográficos, son factores que pueden contribuir al establecimiento de colaboraciones (Leta y Chaimovich, 2002). No obstante, esta colaboración también es incentivada por los propios gobiernos. En España, diversas iniciativas dentro del Programa Ingenio 2010 estaban orientadas a aumentar la colaboración en el sistema español de I+D+i, para reducir la fragmentación y aumentar la masa crítica investigadora (Ministerio de Ciencia e Innovación, 2010).

7. NOTAS

- [1] Los OPI incluyen el CSIC, el Instituto Nacional de Técnica Aeroespacial (INTA), el Instituto de Salud Carlos III (ISCIII), el Instituto Geológico y Minero de España (IGME), el Instituto Español de Oceanografía (IEO), el Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), el Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), y el Instituto de Astrofísica de Canarias (IAC).
- [2] Agricultura/Biología/Medio Ambiente, Biomedicina, Ciencias Sociales, Física, Humanidades, Ingeniería/Tecnología, Matemáticas, Medicina Clínica, Multidisciplinar y Química. El área Multidisciplinar

En resumen, durante el periodo 2000–2011 se pone de manifiesto que todos los sectores experimentan un aumento en su producción, aunque destacan las Entidades sin Ánimo de Lucro, y en especial las EISAL, que con una importante actividad en Biomedicina, multiplican por 12 su producción. Estas entidades presentan grandes equipos; muestran, junto a los OPI, los mayores valores de impacto relativo en las distintas áreas temáticas; y contribuyen por encima de lo esperado al núcleo del 1% de publicaciones españolas más citadas, de lo que se deduce su importancia en el panorama actual de la investigación en el país. La Universidad, responsable de dos terceras partes de la producción nacional, desempeña un papel nuclear en la investigación española y es el único sector especializado en determinados campos como Ciencias Sociales, Humanidades o Matemáticas. La participación del sector Empresas es limitada y, aunque tiende a aumentar el impacto y número de su publicaciones, se requieren todavía esfuerzos para incrementar la implicación del sector en la investigación y su colaboración con el sector académico.

6. AGRADECIMIENTOS

Se agradece la financiación del CSIC (proyecto 200410E605), que ha sido esencial para poder llevar a cabo esta investigación. Igualmente queremos agradecer los comentarios de los revisores que han contribuido a mejorar el trabajo.

ACKNOWLEDGEMENTS

We want to thank the funding of CSIC (research project 200410E605) which has been essential for the development of the research. We are very grateful to the referees who have contributed to improve the article with their comments.

- incluye revistas que publican artículos sobre distintas disciplinas, como *Nature* o *Science*.
- [3] Datos obtenidos mediante consulta en WoS, 18/06/2014.
- [4] Entre las instituciones de la Administración con alta actividad se pueden señalar los jardines botánicos y museos municipales (administración local), las consejerías y organismos dependientes de las mismas (administración regional) y los ministerios e instituciones dependientes como el Instituto Geográfico Nacional o el Banco de España (administración central).
- [5] Entre los centros que contribuyen a explicar la especialización temática del sector Administra-

ción Pública en Agricultura, Física y Medicina Clínica se pueden mencionar: el Instituto Valenciano de Investigaciones Agrarias (IVIA) y el Instituto Andaluz de Investigación Agraria y Pesquera (IFAPA) (Agricultura/Biología/Medio Ambiente); instituciones dependientes de ministerios como la Agencia Estatal de Meteorología y el Instituto Geográfico Nacional (Física); e instituciones dependientes de los gobiernos regionales, como las consejerías de Sanidad (Medicina Clínica).

- [6] El elevado número medio de autores de la producción de los OPI se explica por su especializa-

ción en Física (*big science*), a través de centros como el Instituto Astrofísico de Canarias o el Instituto Español de Oceanografía.

- [7] El mayor uso de las revistas españolas se detecta en el Sector Sanitario (21% de su producción), seguido por la Administración y las Empresas (12% cada sector), la Universidad (7%), las ESAL (6%) y los OPI (3%). Aunque algunas revistas están escritas en inglés, el porcentaje de artículos en español muestra cifras muy similares: máximo en el sector sanitario (19% de su producción), seguido por la Administración (12%), y mínimo en ESAL (5%) y OPI (2%).

8. REFERENCIAS

- Abramo, G.; D'Ángelo, C.A.; Di Costa, F.; Solazzi, M. (2009). University-industry collaboration in Italy: A bibliometric examination. *Technovation*, 29 (6-7), 498-507. <http://dx.doi.org/10.1016/j.technovation.2008.11.003>
- Adams, J. (2012). Collaborations: The rise of research networks. *Nature*, 490 (7420), 335-336. <http://dx.doi.org/10.1038/490335a> / PMID:23075965
- Barlan, J.R. (2010). *Measuring economic inequalities: Lorenz curve, coefficient of variation and Gini coefficient*. Available at: https://docs.google.com/document/d/1yBD5K5k3AesA_uf9uPn3-FHU8O1qi8eVnDbsWsdKueQ/edit
- Batagelj, V.; Mrvar, A. (2014) *Pajek: Program for Analysis and Visualization of Large Networks, Reference Manual, V 3.15*. http://dx.doi.org/10.1007/978-1-4614-6170-8_310
- Bordons, M. (2004). Hacia el reconocimiento internacional de las publicaciones científicas españolas. *Revista Española de Cardiología*, 57 (9), 799-802. [http://dx.doi.org/10.1016/S0300-8932\(04\)77198-5](http://dx.doi.org/10.1016/S0300-8932(04)77198-5)
- Bordons, M.; Barrigón, S. (1992). Bibliometric Analysis of publications of Spanish pharmacologists in the SCI (1984-1989). Part II. Contribution to subfields other than Pharmacology and Pharmacy. *Scientometrics*, 25 (3), 425-446. <http://dx.doi.org/10.1007/BF02016930>
- Bordons, M.; Morillo, F.; Gómez, I.; Moreno, L.; Aparicio, J.; González-Albo, B. (2014). *La actividad científica del CSIC a través de indicadores bibliométricos: (Web of Science, 2009-2013)*. Madrid: IFS-UTAI-CCHS-CSIC. PMID:24799194
- Bordons, M.; Sancho, R.; Morillo, F.; Gómez, I. (2010). Perfil de actividad científica de las universidades españolas en cuatro áreas temáticas: un enfoque multifactorial. *Revista Española de Documentación Científica*, 33 (1), 9-33. <http://dx.doi.org/10.3989/redc.2010.1.718>
- Chang, Y.W. (2014) Exploring scientific articles contributed by industries in Taiwan. *Scientometrics*, 99 (2), 599-613. <http://dx.doi.org/10.1007/s11192-013-1222-2>
- Comisión Interministerial de Ciencia y Tecnología (2007). *Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2008-2011*. Madrid: Fundación Española para la Ciencia y la Tecnología (FECYT).
- Cruz-Castro, L.; Sanz-Menéndez, L.; Martínez, C. (2012). Research centers in transition: patterns of convergence and diversity. *Journal of Technology Transfer*, 37 (1), 18-42. <http://dx.doi.org/10.1007/s10961-010-9168-5>
- De Filippo, D.; Lascurain, M.L.; Casado, E. (2013). Patrones de actividad científica de los diversos sectores institucionales en el Sistema Español de Ciencia, Tecnología e Innovación. Madrid: *Actas del XI Congreso de Sociología*, Vol. II, 1002-1013.
- Díaz-Faes, A.A.; González-Albo, B.; Galindo, M.P.; Bordons, M. (2013). HJ-Biplot como herramienta de inspección de matrices de datos bibliométricos. *Revista Española de Documentación Científica*, 36 (1), 1-16. <http://dx.doi.org/10.3989/redc.2013.1.988>
- Glänzel, W.; Schubert, A. (2001). Double effort = double impact? A critical view at international co-authorship in chemistry. *Scientometrics*, 50 (2), 199-214. <http://dx.doi.org/10.1023/A:1010561321723>
- Godin, B.; Dore, C.; Larivière, V. (2002). The production of knowledge in Canada: Consolidation and diversification. *Journal of Canadian Studies- Revue D'Etudes Canadiennes*, 37 (3), 56-70.
- Godin, B.; Gingras, Y. (2000). The place of universities in the system of knowledge production. *Research Policy*, 29 (2), 273-278. [http://dx.doi.org/10.1016/S0048-7333\(99\)00065-7](http://dx.doi.org/10.1016/S0048-7333(99)00065-7)
- Gómez, I.; Bordons, M.; Morillo, F.; Moreno, L.; Aparicio, J.; Candelario, A.; González-Albo, B.; Herrero, M. (2009). *Indicadores de Producción Científica y Tecnológica de la Comunidad de Madrid (2004-2008)*. Madrid: IEDCYT.
- Grupo Bibliometría (2006). *La investigación del CSIC a través de sus publicaciones científicas de difusión internacional (1981-2003)*. Madrid: CINDOC, CSIC.

- Grupo Scimago (2007). La productividad ISI de las Universidades españolas. *El Profesional de la Información*, 16(4), 354–358. <http://dx.doi.org/10.3145/epi.2007.jul.11>
- Hackett, E. (2005). Introduction: special guest-edited issue on scientific collaboration. *Social Studies of Science*, 35 (5), 667–671. <http://dx.doi.org/10.1177/0306312705057569>
- Huang, M.H.; Chang, H.W.; Chen, D.Z. (2006). Research evaluation of research-oriented universities in Taiwan from 1993 to 2003. *Scientometrics*, 67 (3), 419–435. <http://dx.doi.org/10.1556/Scient.67.2006.3.6>
- Kamada, T.; Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31 (1), 7–15. [http://dx.doi.org/10.1016/0020-0190\(89\)90102-6](http://dx.doi.org/10.1016/0020-0190(89)90102-6)
- Katz, J. S.; Hicks, D. M. (1996). A systemic view of British science. *Scientometrics*, 35 (1), 133–154. <http://dx.doi.org/10.1007/BF02018237>
- Lander, B. (2013). Sectoral collaboration in biomedical research and development. *Scientometrics*, 94 (1), 343–357. <http://dx.doi.org/10.1007/s11192-012-0776-8>
- Lehvo, A.; Nuutinen, A. (2006). *Finnish science in international comparison*. Helsinki: Academy of Finland (Publications of the Academy of Finland; 15/2006).
- Leta, J.; Chaimovich, H. (2002). Recognition and international collaboration: the Brazilian case. *Scientometrics*, 53 (3), 325–335. <http://dx.doi.org/10.1023/A:1014868928349>
- Maltrás, B.; Quintanilla, M.A. (1992). *La producción científica española: 1981–89*. Madrid: Consejo Superior de Investigaciones Científicas.
- Ministerio de Ciencia e Innovación (2010). *Programa Ingenio 2010: Balance de actuaciones*. http://www.idi.mineco.gob.es/stfls/MICINN/Prensa/FI-CHEROS/2010/Versi%C3%B3n_Informe_Ingenio%202010_para_prensa.pdf.
- Moed, H.F. (2005). *Citation analysis in research evaluation*. Amsterdam: Springer.
- Morillo, F.; Aparicio, J.; González-Albo, B.; Moreno, L. (2013). Towards the automation of address identification. *Scientometrics*, 94 (1), 207–224. <http://dx.doi.org/10.1007/s11192-012-0733-6>
- Moya-Anegón, F. de; Solís Cabrera, F.M.; Corera-Álvarez, E.; Chinchilla-Rodríguez, Z.; Gómez-Crisóstomo, R.; González-Molina, A.; Vargas-Quesada, B. (2008). *Indicadores bibliométricos de la producción científica de Andalucía 2003–2005*. Sevilla: Junta de Andalucía, Consejería de Innovación, Ciencia y Empresa.
- Moya-Anegón, F. de; Chinchilla-Rodríguez, Z.; Corera-Álvarez, E.; González-Molina, A.; López-Illescas, C.; Vargas-Quesada, B. (2013). *Indicadores bibliométricos de la actividad científica española 2010*. Madrid: FECYT.
- Moya-Anegón, F. de; López-Illescas, C.; Moed, H.F. (2014). How to interpret the position of private sector institutions in bibliometric rankings of research institutions. *Scientometrics*, 98 (1), 283–298. <http://dx.doi.org/10.1007/s11192-013-1087-4>
- National Science Board (2014). *Science & Engineering Indicators*. Arlington VA: National Science Foundation (NSB 14–01).
- OCDE (2003). *Manual de Frascati, 2002: Medición de las actividades científicas y tecnológicas: Propuesta de Norma Práctica para Encuestas de Investigación y Desarrollo Experimental*. Madrid: FECYT.
- Ortega, J. L.; Aguillo, I. F. (2010). Shaping the European research collaboration in the 6th Framework Programme health thematic area through network analysis. *Scientometrics*, 85 (1), 377–386. <http://dx.doi.org/10.1007/s11192-010-0218-4>
- Rovira, L.; Méndez-Vásquez, R.I.; Suñén-Pinyol, E.; Camí, J. (2007). *Caracterització bibliomètrica de la producció científica a Catalunya, 1996–2006. Informe*. Barcelona: AGAUR-PRBB.
- Science and Technology Indicator Project Team, National Institute of Science and Technology Policy (NISTEP), Ministry of Education, Culture, Sports, Science and Technology (MEXT) (2004). *Science and technology indicators: 2004. A systematic analysis of Science and Technology Activities in Japan*. Japón: MEXT.
- Sort García, M.L.; Rodríguez Méndez, C. (2012) *Análisis ICONO: SCImago Institutions Ranking (SIR) 2012: Documento de Trabajo 3/2012*. Madrid: FECYT. PMID:22258924
- Tijssen, R.; Hollanders, H.; Kanerva, M. (2010). *Science and Technology Indicators 2010: summary report on the Netherlands*. Leiden: Netherlands Observatory of Science and Technology (NOWT).
- Van Leeuwen, T.N.; Moed, H.F.; Tijssen, R.J.W.; Visser, M.S.; van Raan, A.F.J. (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics*, 51 (1), 335–346. <http://dx.doi.org/10.1023/A:1010549719484>
- Wuchty, S.; Jones, B.F.; Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science* 316, 1036–1039. <http://dx.doi.org/10.1126/science.1136099> / PMID:17431139



ESTUDIOS / RESEARCH STUDIES

**Análisis de la producción científica hispana en mejora continua:
1990-2011**

Lidia Sánchez*, Beatriz Blanco*

*Universidad de Cantabria. Departamento de Administración de Empresas.
Facultad de Ciencias Económicas y Empresariales. Santander. Cantabria.
Correo-e: sanchezrl@unican.es, blancob@unican.es

Recibido: 11-02-2015; 2ª versión: 18-05-2015; Aceptado: 15-06-2015.

Cómo citar este artículo / Citation: Sánchez, L. y Blanco, B. (2016). Análisis de la producción científica hispana en mejora continua: 1990-2011. *Revista Española de Documentación Científica*, 39(1): e116. doi: <http://dx.doi.org/10.3989/redc.2016.1.1264>

Resumen: El objetivo de este trabajo es realizar una revisión del campo de la Mejora Continua en la literatura. Para ello se realiza una revisión rigurosa tanto con métodos cualitativos como cuantitativos a partir de los trabajos recogidos en la base de datos Dialnet. Los principales resultados muestran que la mejora continua es un campo importante en la literatura hispana, aunque debe potenciarse en ciertas temáticas. También se detectó una clara orientación divulgativa del tema resaltando así la importancia de la mejora continua para los profesionales del sector. El trabajo tiene dos aportaciones principales. En primer lugar, es un estudio novedoso que no se había realizado previamente en la literatura hispana. Y, en segundo lugar, matiza y complementa las conclusiones obtenidas por revisiones internacionales según las cuales la mejora continua parecía no estar desarrollada en la literatura hispana.

Palabras clave: Mejora continua; kaizen; revisión bibliográfica; temáticas; evolución; literatura hispana; técnicas bibliométricas; metodologías.

Review of professional literature from the Hispanic world on the field of continuous improvement: 1990-2011

Abstract: The aim of this paper is to present the results of a review of Hispanic literature about continuous improvement. A rigorous review of papers included in the Dialnet data base was conducted, combining both qualitative and quantitative methods. The main results show that continuous improvement is an important field in Hispanic output although certain topics should be delved into more deeply. General knowledge journals aimed at practitioners are those that tend to publish papers related to continuous improvement, thus highlighting the relevance of the topic for practitioners. The paper offers two main contributions. First, it is novel research not been previously done in the Hispanic context. Secondly, it clarifies and complements the conclusions drawn by international reviews according to which continuous improvement seemed to be scarcely addressed in Hispanic output.

Keywords: Continuous improvement; kaizen; literature review; topics; evolution; Hispanic literature; bibliometric techniques; methodologies.

Copyright: © 2016 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0.

1. INTRODUCCIÓN

En la actualidad las empresas desarrollan su actividad en un entorno extremadamente competitivo y cambiante. En este contexto, la mejora continua es considerada como un elemento clave para conseguir la excelencia empresarial, tal y como expresaron De Leede y Kees (1999).

El término mejora continua tiene su origen en la expresión Kaizen, palabra compuesta de origen japonés que engloba dos conceptos: Kai (cambio) y Zen (para mejorar) (Singh y Singh, 2009).

La mejora continua puede definirse como aquel proceso constante de mejora incremental que se realiza en la empresa con la participación de todo el personal. En el trabajo desarrollado por Sanchez y Blanco (2014) puede consultarse un amplio listado de definiciones del concepto.

Las razones que, durante los últimos años, han impulsado la mejora continua son de muy diversa índole. En nuestra opinión, los tres principales motivos que han propiciado el crecimiento y expansión del movimiento de la mejora continua son: los cambios en el entorno empresarial, la aparición de nuevos sistemas de gestión y la relevancia indudable de la gestión de la calidad.

En este contexto, la mejora continua adquiere vital importancia dado que las empresas no deben estancarse, sino que deben ser flexibles, adaptarse y mejorar continuamente. Todo ello orientado a satisfacer de la manera más eficiente posible los requerimientos cambiantes de los clientes, en un entorno también en constante cambio.

La importancia de la mejora continua en el mundo real también ha tenido su reflejo en el mundo académico (Serrano Bedia y otros, 2007) y, como veremos ampliamente en apartados posteriores, el tema de la mejora continua ha sido tratado desde múltiples perspectivas.

Precisamente, dada la multitud de estudios sobre mejora continua, parece oportuno realizar un análisis de la literatura existente hasta el momento. Este tipo de revisiones son útiles de cara a establecer las agendas de investigación en las que se identifican los avances que se han producido hasta el momento, y se plantean las líneas futuras que aún están por desarrollar (Correa i Domènech y otros, 2009), siendo de especial interés para autores y editores (González Alcaide y otros, 2008).

En la literatura existen diversos trabajos que revisan el concepto teórico de la mejora continua desde múltiples enfoques (Dahlggaard-Park y otros, 2013; Sanchez y Blanco, 2014; Singh y

Singh, 2009; Singh y Singh, 2012; Suarez-Barraza y Angel Miguel-Davila, 2011; Suárez-Barraza, 2008). En todos los casos, los trabajos antes citados hacen uso de bases de datos internacionales para identificar y analizar cuáles son las principales temáticas y tendencias del campo de la mejora continua. Si bien es cierto que el uso de bases internacionales tiene ventajas indudables, como el hecho de obtener una visión general de la evolución global de la mejora continua, presenta un problema a tener en cuenta, el posible sesgo hacia trabajos escritos en inglés. El hecho de trabajar con bases de datos internacionales, que incluyen en su mayoría publicaciones en lengua inglesa, plantea la posibilidad de que haya un bloque de literatura en otros idiomas que no se está teniendo en consideración.

Este trabajo pretende dar respuesta a este problema mediante la realización de una revisión objetiva de la evolución de los artículos científicos publicados en la literatura hispana sobre mejora continua durante el periodo 1990-2011.

Consideramos que este tipo de revisiones más concretas pueden servir de complemento a las revisiones internacionales. A través de ellas, además de solucionar el problema del sesgo antes señalado, se pueden conocer mejor los estudios desarrollados en un determinado ámbito geográfico. Esto último es especialmente interesante cuando, como en este caso, la definición del campo de estudio tiene como objetivo posterior la definición y realización de un trabajo empírico sobre el tema en un área geográfica específica. Para este tipo de estudios es muy importante identificar todos los estudios realizados, incluso si no se encuentran recogidos en bases internacionales.

En el campo concreto de la mejora continua, no hay estudios de revisión en el ámbito hispano. Por ello, consideramos que la presente investigación es novedosa puesto que combinando métodos cualitativos y cuantitativos analiza el estado del arte del concepto en cuestión.

Teniendo esto en consideración, el resto del trabajo se estructura de la siguiente manera. En el siguiente apartado se presentan las fuentes de información consultadas y la metodología seguida durante el análisis que, como podrá observarse, sigue la misma estructura que el trabajo de Sanchez y Blanco (2014) por motivos de claridad y también para poder comparar algunos resultados. En la sección tercera se presentan los resultados y, para terminar, la sección cuarta recoge la discusión y las principales conclusiones del estudio.

2. MATERIALES Y MÉTODOS

Para llevar a cabo la investigación se ha realizado una revisión consistente en el análisis de 275 artículos relacionados con el tema de la mejora continua en la literatura hispana. La utilización del término *hispana* en lugar de *española* se debe a que el estudio pretende ser inclusivo, teniendo en consideración todos aquellos artículos escritos, no sólo por autores de procedencia española, sino también autores de países e instituciones de América Central y América del Sur. El concepto hispano se utiliza, por lo tanto, de acuerdo con la acepción de la Real Academia de la Lengua que lo define como todo aquello perteneciente o relativo a las naciones de hispanoamérica.

Para la identificación de los artículos se ha utilizado la base de datos Dialnet. Ésta es la mayor hemeroteca de artículos científicos hispanos en Internet, aunque también recoge libros, conferencias y tesis doctorales entre otros documentos. Desde el año 2004 el número de usuarios ha crecido exponencialmente pasando de 7.343 a 957.060 usuarios en 2012, mostrando así la relevancia de esta base de datos. Además, en marzo de 2012, superó los 3,5 millones de documentos, estando el 15% de los mismos libremente disponibles.

Si bien es cierto que existen otras bases de datos especializadas en el mundo de la economía y la empresa, como por ejemplo ABI Inform Database, el hecho de que no estén enfocadas en el mundo hispano las hacía menos adecuadas para ser utilizadas como fuente de información en este estudio.

Los criterios de búsqueda en la citada base han sido los siguientes:

- El primer filtro de la búsqueda fue el tipo de documento. Así, sólo se incluyeron los artículos. Se decidió esto porque, en la actualidad, son varios los autores que definen el artículo como la unidad de transmisión de conocimiento por excelencia (Arguimbau-Vivó y otros, 2013; Ramos Rodríguez y Ruiz Navarro, 2008; Ramos, 2004).
- El segundo criterio fue el periodo temporal a abarcar. Se decidió que el periodo temporal analizado abarcaría desde el año 1990 hasta el año 2011. Puesto que el análisis se comenzó a mediados de 2012, este año no se incluyó para ser lo más exactos en los datos.
- El tercer, y último, criterio fue la temática. Como es lógico únicamente estábamos interesados en aquellos artículos que estuvieran relacionados con la mejora continua. Para ello se determinó que los artículos debían contener en su título, abstract o palabras clave las palabras "mejora

continua" o "kaizen". El uso de las palabras clave como método de búsqueda es bastante común y se ha utilizado habitualmente en otras revisiones del campo de la gestión (Bedia y otros, 2013; Correa i Domènech y otros, 2009; Richart-Ramón y otros, 2011). Se introdujeron palabras clave amplias para poder identificar el mayor bloque de literatura posible y evitar la exclusión de trabajos importantes. Tras realizar la búsqueda los trabajos fueron revisados manualmente para comprobar, primero, que no había duplicados y, segundo, que efectivamente trataban el tema de la mejora continua. La muestra final obtenida se compuso de un total de 275 artículos (la lista de referencias de estos artículos no se ha incluido dada su longitud, pero puede solicitarse a los autores).

El análisis realizado combinó resultados cuantitativos y cualitativos. Con el objetivo de ser lo más claros posible, se siguió la estructura planteada y testada ya en otros trabajos según la cual los datos se analizaban desde tres perspectivas (Houy y otros, 2010; Sanchez y Blanco, 2014):

1. Perspectiva bibliométrica: análisis cuantitativo de los resultados que permite "medir" el desarrollo de un determinado campo de estudio teniendo en cuenta, entre otros, criterios temporales o de publicación. Este análisis incluye el análisis del número de contribuciones por año, el número de contribuciones por revista, los autores más prolíficos, así como las instituciones y países de pertenencia de los autores.
2. Perspectiva de contenido: análisis mixto que intenta cuantificar el número de estudios desarrollados en un campo de investigación definiendo temas más específicos dentro del propio campo.
3. Perspectiva metodológica: analiza la metodología aplicada en cada uno de los trabajos primarios objeto de estudio. Además del tipo de metodología utilizado en cada uno de los artículos, se analizan los sectores objeto de estudio.

3. RESULTADOS

3.1. Análisis bibliométrico de la literatura

A. Número de contribuciones por año

La búsqueda realizada abarcaba el periodo comprendido entre 1990 y 2011. Durante dicho periodo se localizaron 275 artículos relacionados con la mejora continua. No obstante, dado que durante los años 1990-1993, no se encontró ningún artículo relacionado con el tema, este periodo se ha omitido en el análisis (Ver Figura 1).

Figura 1. Número de artículos sobre mejora continua

De acuerdo con el gráfico anterior la evolución del número de artículos publicados sobre mejora continua ha seguido una tendencia creciente a largo plazo pero con muchas desigualdades anuales. En los últimos años, nos encontramos en un periodo bajista pero, dada la evolución en etapas anteriores, no podemos concluir que esto signifique un decaimiento en el interés sobre el tema.

B. Número de artículos por revista

Con este análisis se pretendía conocer cuál o cuáles eran las publicaciones más idóneas relacionadas con la mejora continua. Dado el gran número de publicaciones obtenidas en los resultados, a continuación se muestran aquellas que publicaron tres artículos o más relacionados con la mejora continua (Tabla I).

Tabla I. Listado de publicaciones y número de artículos publicados

Revista	Nº Artículos	Porcentaje artículos sobre el total
Qualitas hodie	88	32,00
Forum Calidad	47	17,09
UNE: boletín mensual de AENOR	12	4,36
Capital Humano	8	2,91
Revista de calidad asistencial	5	1,82
Economía industrial	4	1,45
Intangible Capital	3	1,09
Alta Dirección	3	1,09
DYNA	3	1,09
GCG : Revista de Globalización, Competitividad y Gobernabilidad	3	1,09
Mantenimiento: ingeniería industrial y de edificios	3	1,09

La primera observación a realizar es el elevado número de publicaciones existente. En total, 95 publicaciones habían publicado algún artículo relacionado con la mejora continua a lo largo del periodo analizado. Ahora bien, de entre ellas, un total de 69 únicamente habían publicado un artículo y 15 dos artículos. Las revistas con un mayor número de artículos publicados están incluidas en la tabla I.

En segundo lugar se detecta claramente que la publicación *Qualitas Hodie* es, sin duda, la revista que más artículos relacionados con la mejora continua ha publicado. En los siguientes puestos del ranking se sitúan *Forum Calidad* y *UNE (Boletín mensual de AENOR)*. Entre las tres publicaciones acumulan el 50% de los artículos identificados sobre el tema. Llama la atención que las tres revistas antes citadas, no sean revistas científicas sino de divulgación. Este hecho muestra el gran interés existente por la mejora continua entre los profesionales, justificando además la necesidad de continuar investigando en el tema, lo que a su vez favorecerá el acercamiento entre el mundo académico y el profesional.

Además, el análisis de la tabla I nos permite observar que el tema de la mejora continua es multidisciplinar, tal y como demuestra la existencia de revistas con perfil de gestión, de sanidad o de ingeniería.

C. Número de artículos por autor

En los 275 artículos analizados, se identificaron 283 coautores. La tabla II muestra el listado con los 9 investigadores más prolíficos en el tema de la mejora continua. Así, Penacho aparece como el coautor más productivo con 34 artículos publicados.

En la tabla II se hace referencia también a la institución de pertenencia de estos nueve autores. Dos de los autores incluidos no pertenecen a ninguna institución académica. El hecho de que aparezcan entre los autores más prolíficos es otro de los argumentos que refuerzan la idea del interés divulgativo del tema objeto de estudio.

Otro resultado interesante es que tres de los nueve autores pertenecen a la misma universidad, Universidad de Navarra, lo que parece indicarnos que existe un grupo de investigación potente en esta área en la citada universidad.

Con el objetivo de profundizar en este aspecto, se ha procedido a analizar la institución de pertenencia y el país de procedencia de todos los autores. Del total de autores, no se pudo identificar la institución de pertenencia de 29 de ellos; 138 eran profesionales, es decir, no eran profesores de universidad ni académicos sino trabajadores de empresas privadas; y, por último, los otros 116 autores pertenecían a un total de 50 universidades. El gran número de profesionales es un resultado lógico que concuerda con el carácter divulgativo de las revistas encontradas.

En la tabla III se recoge un listado con las instituciones con un mayor número de autores. En concreto, se han incluido las que tienen 3 o más autores trabajando en el tema objeto de estudio. Se observa que en la Universidad Complutense de Madrid, en la Universidad de Navarra y en la Universidad Politécnica de Valencia existen grupos de investigación dedicados a la investigación sobre mejora continua.

En la distribución por países que se recoge en la Tabla IV, España aparece en el primer lugar, a una distancia notable del país situado en la siguiente posición.

Tabla II. Número de contribuciones por coautor

Autor	Nº Artículos	Institución	País
Penacho, J.L.	34	Profesional	-
Suarez-Barraza, M.F.	8	Tecnológico de Monterey	México
Mateo Dueñas, R.	5	Universidad de Navarra	España
Martín Castilla, J.I.	4	Universidad Autónoma de Madrid	España
Santos García, J.	4	Universidad de Navarra	España
González Durán, F.	3	Profesional	-
Jaca, M.C.	3	Universidad de Navarra	España
Marín García, J.A.	3	Universidad Politécnica de Valencia	España

Tabla III. Instituciones firmantes de los artículos

Instituciones firmantes	País	Nº de autores
Universidad Complutense de Madrid	España	8
Universidad de Navarra	España	8
Universidad Politécnica de Valencia	España	8
Universidad de Sevilla	España	6
Universidad de Valladolid	España	5
Universidad Carlos III de Madrid	España	4
Universidad de la Frontera	Chile	4
Universidad de Valencia	España	4
Universidad Politécnica de Tlaxcala	México	4
Universidad Autónoma de Ciudad Juárez	México	3
Universidad Autónoma de Madrid	España	3
Universidad de Barcelona	España	3
Universidad de Cádiz	España	3
Universidad de Cartagena de Indias	Colombia	3
Universidad de Córdoba	España	3
Universidad de Huelva	España	3
Universidad Estácio De Sa	Brasil	3

Tabla IV. Países firmantes de los artículos

País	Nº de autores
España	87
México	9
Colombia	5
Chile	5
Venezuela	2
Brasil	3
Costa Rica	1
Cuba	1
Inglaterra	1
Nicaragua	1

3.2. Análisis de la literatura desde la perspectiva del contenido

El objetivo de este análisis era identificar cuáles eran las grandes temáticas tratadas dentro del campo de la mejora continua. Conocer cuáles son las tendencias en el campo, no sólo permite conocer cuál es el estado actual del tema, sino que facilita el planteamiento de futuras líneas de investigación.

Tras una primera lectura de los resúmenes de los 275 artículos se establecieron por consenso nueve temáticas en las que iban a clasificarse. Estas temáticas son las que se utilizaron en el estudio de Sanchez y Blanco (2014), donde se realiza una revisión internacional sobre el concepto. El hecho de utilizar las mismas temáticas permitiría realizar comparaciones posteriores. Estas nueve temáticas eran las siguientes:

- **Concepto:** bajo este epígrafe se agrupan los estudios orientados al análisis del concepto de "mejora continua" propiamente dicho.
- **Implantación:** bajo este título se agruparían aquellos estudios que presentaran resultados relacionados con la implantación de la mejora continua en una organización. Incluía tanto estudios empíricos en los que se presentaba un caso de estudio, como estudios teóricos donde se proponían metodologías aún por testar.
- **Factores:** engloba investigaciones que analizan los obstáculos, los facilitadores, los beneficios e inconvenientes que se derivan de la implantación de un programa de mejora continua.
- **Herramientas:** incluía trabajos centrados en la aplicación de una o varias herramientas de cualquier tipo como medio para el desarrollo de un programa de mejora continua.

- **Cultura:** todos aquellos estudios relacionados con temas de cultura empresarial, buenas prácticas, responsabilidad social y medioambiental, seguridad en el trabajo...
- **Control:** incluye las investigaciones relacionadas con el establecimiento de mediciones, indicadores... Dada la importancia del establecimiento de sistemas de control a la hora de implantar un sistema de mejora continua, consideramos oportuno dedicarle un apartado independiente a este tema en vez de incluirlo en el epígrafe de implantación, donde también hubiera tenido cabida.
- **Filosofías de gestión:** agrupa artículos donde la mejora continua se relaciona con otras filosofías y sistemas de gestión como son el lean management, la gestión por procesos, la gestión de la calidad total o la teoría de las limitaciones.
- **Innovación:** engloba estudios en los que se establecen relaciones entre la mejora continua y la innovación.
- **Recursos Humanos:** en este tema se agrupan los estudios que analizan las prácticas de recursos humanos (formación, motivación, participación...) con el éxito de la mejora continua y viceversa. Es decir, la influencia de la mejora continua en los recursos humanos de la organización (satisfacción de los empleados).

Tras fijar las temáticas se revisaron los artículos y se los clasificó en un máximo de dos temáticas. Esto hace que los resultados aquí presentados no sean en términos absolutos por lo que han de ser interpretados con cautela. No obstante, consideramos que los resultados permiten hacerse una idea de la evolución del campo de investigación, así como establecer futuras líneas de investigación. En la tabla V se resumen los resultados obtenidos.

Tabla V. Temas tratados en los estudios de mejora continua

Temáticas	Porcentaje
Implantación	32%
Herramientas	26%
Filosofías de gestión	14%
Concepto	9%
RRHH	6%
Control	5%
Cultura	4%
Factores	2%
Innovación	1%

La temática más común entre los artículos analizados es la de "Implantación". Esta temática incluye aquellos trabajos en los que se analiza un caso concreto de implantación de la mejora continua, además de aquéllos en los que se definen metodologías de implantación. Este tema guarda una estrecha relación con el siguiente ya que en muchos casos las implantaciones incluyen el uso o análisis de determinadas herramientas.

En segundo lugar se encuentra la temática de "Herramientas", es decir, abundan los trabajos en los que se analiza el uso de determinadas herramientas para la implantación o durante el desarrollo de un programa de mejora continua. Ante la importancia del tema, se decidió profundizar en el mismo y determinar cuáles eran las herramientas más comunes en los estudios de mejora continua.

La herramienta más importante es la acreditación. A pesar de que la acreditación en sí misma no es una herramienta, sino un reconocimiento o una guía, consideramos que puede ser considerada como tal ya que ayuda al autodiagnóstico y a la mejora continua. En este apartado se identificaron un total de 21 publicaciones que se referían a los sistemas ISO 9000/9001; otras 21 publicaciones adicionales trataban sobre el modelo EFQM de excelencia europeo; y, por último, cuatro de las publicaciones analizadas hacían referencia a otros premios nacionales o internacionales de menor reconocimiento.

Después de la acreditación, las herramientas más utilizadas son la técnica Six Sigma, las tecnologías de la información, el análisis modal de fallos y efectos, y el benchmarking.

Volviendo a las temáticas tratadas, la tercera posición la ocupa la temática de "Filosofías de gestión". Este resultado es esperado ya que en muchas ocasiones la implantación de la mejora continua está asociada a la instauración de un sistema de gestión mayor como puede ser el lean management o la gestión de la calidad total.

El cuarto lugar lo ocupa el tema "concepto", seguido por las temáticas de "Recursos humanos" y "Control". La posición intermedia de recursos humanos es lógica si tenemos en cuenta la importancia de la participación de personal en los programas de mejora continua. Así, la participación de toda la empresa, la motivación del personal o una correcta política de formación para los empleados son sólo algunas prácticas de necesaria implantación para desarrollar un programa de mejora continua (Berling, 2000; Cooney y Sohal, 2004; Higuera-López, 2011; Jun y otros, 2004; Rapp y Eklund, 2007; Warwood y Roberts, 2004).

Los puestos restantes los ocupan, por orden de importancia, la cultura, los factores y la innovación. Estos últimos temas, dado su escaso desarrollo, se presentan como unas futuras líneas de investigación a desarrollar.

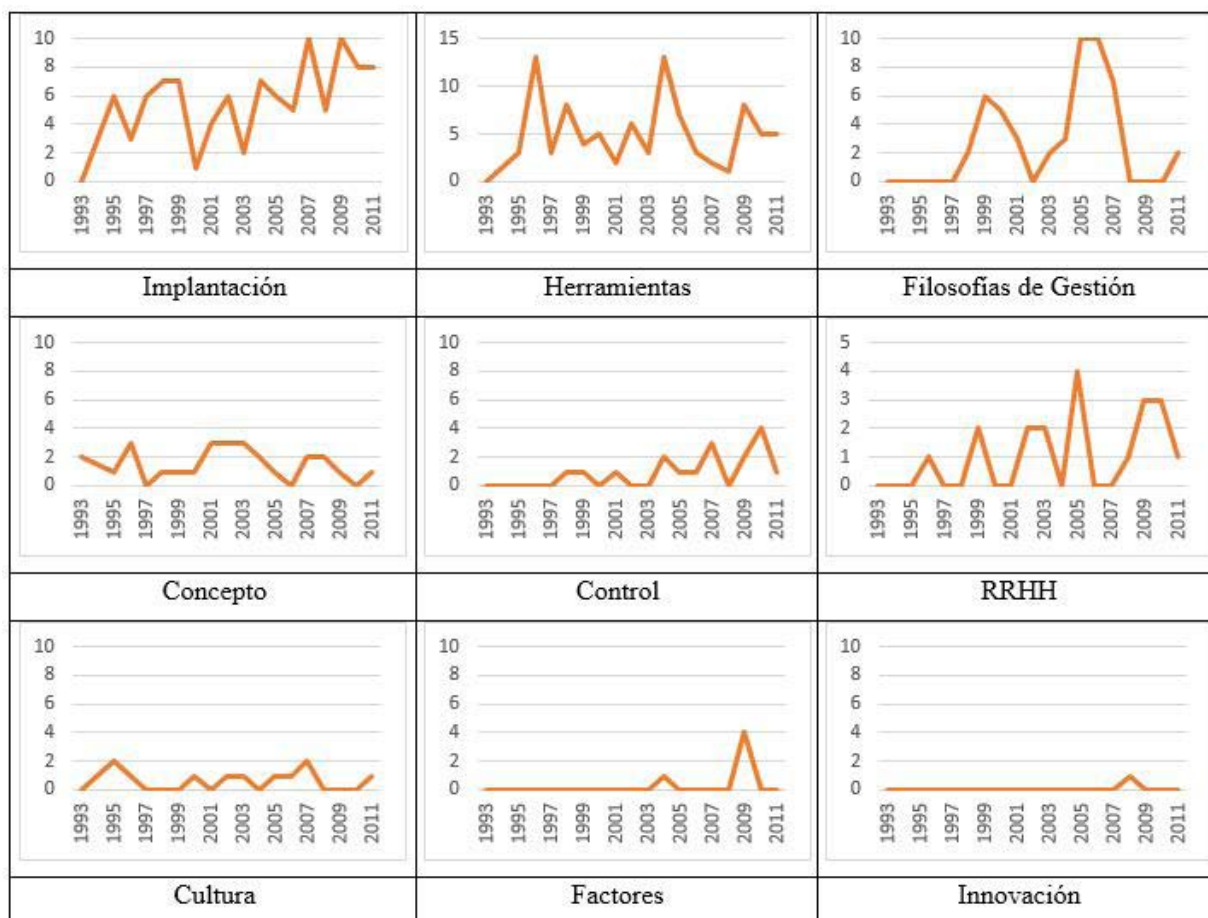
Como análisis complementario se ha procedido a combinar el análisis temático con la evolución temporal de las publicaciones, todo ello con el objetivo de detectar posibles tendencias en las temáticas (Figura 2). A lo largo de los años objeto de estudio, únicamente dos temáticas, implantación y herramientas, mantienen cierta continuidad. Es decir, son las únicas temáticas de las cuales se han publicado artículos a lo largo de todos los años. Así, a pesar de que existen variaciones entre un año y otro, en ambos casos existe una media de 5 artículos publicados por temática. De hecho, es también destacable que, a diferencia de lo que ocurre a nivel general (ver Figura 1), estas dos temáticas no reducen el número de publicaciones a partir del 2008. Este

hecho puede ser indicativo de que la reducción detectada a partir del año 2008, no supone un descenso en el interés por la mejora continua, sino una reducción de las publicaciones en ciertas temáticas. Dicho de otro modo, las temáticas de herramientas e implantación continúan siendo interesantes y el número de publicaciones continúa siendo regular.

En el lado opuesto están las temáticas de innovación y factores. En el caso de la temática de innovación, únicamente aparece un artículo sobre este tema en el año 2008. Por otro lado, sólo hay dos años en los que se han identificado artículos referidos a la temática de los factores.

Por último, en lo referido al resto de temáticas, todas ellas presentan grandes variaciones a lo largo de los diferentes años. No se observa por lo tanto ninguna tendencia ni comportamiento digno de mención que pueda orientar futuras investigaciones en el campo.

Figura 2. Evolución de las temáticas de las publicaciones



3.3. Análisis de la literatura desde la perspectiva metodológica

A. Metodologías de análisis utilizadas

A lo largo de los 20 años analizados aparecen 91 artículos empíricos. Esto es otra muestra del interés práctico y real que existe sobre la mejora continua, hecho que también quedó reflejado en la orientación divulgativa de las revistas más prolíficas en el tema.

Los artículos empíricos son de muy diversa naturaleza y utilizan diferentes metodologías. El objetivo de este apartado consiste en determinar qué metodologías han sido utilizadas en el campo de investigación de la mejora continua.

Tras el análisis, se han identificado tres metodologías. La más común es el caso de estudio, utilizada por 77 artículos. La encuesta se sitúa en un segundo lugar y a una distancia notable de la primera opción, ya que sólo es utilizada por 11 publicaciones. Por último, sólo tres publicaciones hacen uso de la metodología del *action research*.

El hecho de que el caso de estudio sea la metodología más utilizada, concuerda perfectamente con el hecho de que las dos temáticas más tratadas sean "implantación" y "herramientas", ya que en muchos casos se trata de experiencias concretas. Además, dado el grado de madurez del campo de la mejora continua, es lógico pensar que el caso de estudio y la encuesta sean las metodologías más habituales ya que el primero de ellos es común en las primeras fases de desarrollo de un campo de investigación y la segunda es comúnmente utilizada en fases más avanzadas para la comprobación de hipótesis.

B. Sectores analizados en los artículos

En lo referente a los sectores de aplicación de los estudios empíricos, los resultados muestran que un total de 33 artículos empíricos se han aplicado al sector manufacturero, 52 al sector servicios y 6 artículos incluyen empresas de ambos sectores. Sorprende que el número de artículos del sector servicios sea superior al del sector manufacturero, ya que, por lo general el sector manufacturero suele ser pionero.

3.4. Relaciones de dependencia: temáticas, sectores y metodologías

En esta sección se ha procedido a comprobar si existe alguna relación entre tres de las variables antes analizadas: temáticas, sectores y metodologías. Para ello, y con el objetivo de seleccionar la prueba estadística adecuada, en primer lugar se ha

procedido a comprobar si los datos se distribuyen normalmente. En concreto se ha utilizado la prueba Kolmogorov-Smirnov. Los resultados han mostrado que ninguna de las variables se distribuye normalmente. Como consecuencia de lo anterior, se han utilizado pruebas no paramétricas para comprobar si existe o no relación entre las variables.

En primer lugar se comprobó si existe relación entre el sector objeto de estudio y la metodología de investigación utilizada. El resultado fue positivo, es decir, dependiendo del sector en el que se centre el artículo, la metodología usada es diferente. En concreto en los sectores de manufactura, servicios y sanidad predomina el caso de estudio; mientras que en aquellos estudios de sector mixto predomina la encuesta.

En segundo lugar, se repitió el análisis para las variables sector y temática. De nuevo los resultados de la prueba Chi Cuadrado muestran que existe relación entre ambas variables. Así, el interés de las temáticas varía en función del sector actividad.

Por último, se comprobó que sí existe relación entre la temática y la metodología. De este modo, parece que en función de la temática objeto de estudio, el investigador selecciona una u otra metodología de análisis.

4. CONCLUSIONES Y DISCUSIÓN

Este estudio ofrece una revisión del estado del arte del campo de la mejora continua en la literatura hispana combinando análisis cuantitativos y cualitativos.

Su aportación a la literatura es doble. Desde el punto de vista bibliográfico, utiliza una estructura de análisis que combina tres perspectivas: la perspectiva bibliométrica, la perspectiva de contenido y la perspectiva metodológica. Se muestra que es una estructura útil de cara a analizar un campo de investigación. Desde el punto de vista conceptual, el de la mejora continua, el estudio resulta novedoso al no existir otros estudios que, de forma rigurosa, analicen la literatura sobre mejora continua en el ámbito hispano.

En lo referente al posible sesgo existente en las revisiones internacionales hacia el idioma inglés, a continuación exponemos las diferencias y similitudes encontradas entre este estudio y el resto de revisiones identificadas en el apartado introductorio.

Si bien es cierto que los resultados y conclusiones alcanzados en ambas revisiones, como comentaremos a continuación, son bastante similares, parece que las revisiones más concretas, como el presente trabajo, son un complemento deseable

a las revisiones internacionales. Así, a través de esta revisión se ha podido detectar un bloque de literatura hispana mucho mayor al inicialmente detectado por el estudio de Sanchez y Blanco (2014). En ese estudio, se realizaba una revisión de la literatura a nivel internacional haciendo uso de la base de datos Web of Science (WoS) siguiendo la misma metodología que en el presente estudio. En concreto se analizaron 1090 artículos sobre mejora continua publicados entre los años 1980 y 2012. Uno de los resultados más destacables, en relación al desarrollo de la temática por ámbitos geográficos, es el hecho de que España contaba con 50 publicaciones sobre la temática objeto de estudio, situándose en un tercer puesto, pero muy alejada de Estados Unidos y Reino Unido que tenían un total de 422 y 174 artículos respectivamente. Por otro lado, otros países hispanos como México, Argentina, Chile o Venezuela aparecían con 9, 3, 2 y 2 artículos respectivamente. Se concluía, por lo tanto, que el desarrollo de la temática en estos países era muy escaso. Sin embargo, los resultados de este estudio muestran que, pese a no alcanzar las cifras de los países punteros en el tema, la producción hispana sobre mejora continua es más numerosa que lo identificado inicialmente por la revisión internacional. Se demuestra, por tanto, que las revisiones internacionales, aunque útiles, han de completarse con revisiones más específicas.

Este argumento se ve reforzado además por el hecho de que sólo cuatro de los 275 artículos analizados se encontraban en la base de datos WoS. Es decir, el solapamiento entre ambas bases de datos es prácticamente inexistente, y la aportación de la base de datos hispana Dialnet sobre la internacional es significativo.

En cuanto a los resultados de la revisión propiamente dicha, se comentan en las siguientes líneas. En lo que se refiere a la evolución temporal, se ha detectado una tendencia creciente a largo plazo aunque anualmente las variaciones son muy llamativas. Estos fuertes contrastes entre periodos son también detectados en el trabajo de Dajlgaard-Park y otros (2013). Estos autores plantean la posibilidad de que las variaciones anuales en el número de artículos puedan explicarse, al menos parcialmente, en función del mayor o menor grado de interés de los profesionales (que muchas veces depende únicamente de si la temática está o no de moda). Esto es, de si la mejora continua es una moda o no. En los últimos años objeto de análisis se observa una caída; sin embargo, consideramos que la tendencia seguirá siendo creciente y el tema de la mejora continua permanecerá siendo de interés para investigadores y profesionales, conclusión también obtenida por Dahlgaard-Park y otros (2013).

Otra conclusión interesante es, precisamente, el elevado interés que existe hacia la mejora continua desde el mundo profesional. Muestra de ello son la multitud de estudios empíricos que se han encontrado, así como la predominancia de revistas de tipo divulgativo orientadas a profesionales, junto con el elevado número de autores profesionales que se identificaron en el texto.

Del análisis de contenido se obtienen también interesantes conclusiones. Los autores, en primer lugar, definieron nueve temáticas clave (concepto, implantación, factores, metodologías, cultura, control, filosofías de gestión, innovación y recursos humanos). Entre ellas las más repetidas fueron la "implantación" y las "herramientas". En el estudio de Sanchez y Blanco (2014) también se identificaron estas dos temáticas como las más importantes. Por su parte, Dahlgaard-Park y otros (2013) también detectaron que los artículos sobre herramientas y su aplicación eran más comunes y seguían una tendencia creciente a lo largo del tiempo. Estos autores también destacaron la necesidad de evolucionar del estudio del caso particular hacia estudios más teóricos y conceptuales que intentaran describir la mejora continua desde una perspectiva más general.

El tema de "Filosofías de gestión" ocupa la tercera posición por orden de importancia. Ya en la parte introductoria se expresó la estrecha relación que existía entre la mejora continua y las modernas filosofías de gestión, luego no es de extrañar que ésta sea una de las temáticas más comunes en el campo de investigación objeto de análisis. De hecho, Suárez Barraza (2008) ya identificó la mejora continua como un elemento clave de otras filosofías de gestión como la Gestión de la Calidad Total, analizando en profundidad el cuerpo de literatura que lo definía como tal.

La temática de recursos humanos ocupa una posición intermedia. Sin embargo, dada la importancia del personal en cualquier proceso de cambio, cabría esperar un mayor número de estudios en este apartado que proporcionen nuevas perspectivas sobre la influencia de las políticas de recursos humanos en los procesos de implantación de la mejora continua y viceversa.

Por último, existe un grupo de temas (cultura, factores e innovación) cuyo desarrollo es aún incipiente y, por lo tanto, representan futuras líneas de investigación. En lo referente al tema "cultura", el análisis realizado en este estudio muestra que es una de las temáticas menos desarrolladas a lo largo del periodo analizado, coincidiendo con las conclusiones de Singh y Singh (2012). Por otro lado, en el caso de la temática de los factores, Suarez-

Barraza (2008) ya destacó la necesidad de hacer mayor hincapié en los factores críticos de éxito de la mejora continua como un medio para comprender mejor esta filosofía.

En el último apartado del estudio se analizaron las metodologías utilizadas en los estudios empíricos encontrados en la revisión. El caso de estudio ocupa el primer lugar seguido por la encuesta y el *action research*. Este resultado es muestra del estado de madurez del campo de la mejora continua, ya que por lo general en las fases iniciales de carácter exploratorio se hace uso de la metodología del caso y a medida que se avanza se utiliza la encuesta, un método adecuado para la comprobación y testeo de hipótesis. Este resultado es coincidente con lo expresado por Singh y Singh (2009) que entendían el caso de estudio y la encuesta como las principales metodologías de estudio en el campo de la mejora continua.

En base a todo lo anterior, y pese a las similitudes en los resultados, se podría concluir que es oportuno complementar las revisiones internacionales con revisiones centradas en el análisis de bases de datos de cobertura local. Esto permitiría detectar nichos o características concretas propias

de un bloque concreto, en este caso la literatura hispana. Además de las diferencias detectadas en el contenido, también se han detectado diferencias a nivel cuantitativo, identificándose un mayor número de publicaciones en la revisión nacional lo que, sin duda, enriquece la revisión basada en bases internacionales.

Dada la rápida evolución de la investigación, consideramos oportuno y necesario realizar de forma periódica este tipo de estudios. Así, no sólo podrá definirse el estado del arte del campo en cuestión, sino que podrán establecerse las líneas futuras de investigación de forma más adecuada.

5. AGRADECIMIENTOS

Los autores quieren agradecer al Ministerio de Educación español por financiar esta investigación a través del programa de Formación de Profesorado Universitario (FPU).

ACKNOWLEDGEMENTS

This work was supported by the Spanish Ministry of Education under FPU Grant.

6. REFERENCIAS

- Arguimbau-Vivó, L.; Fuentes-Pujol, E.; Gallifa-Calatayud, M. (2013). Una década de investigación documental sobre ciencia de la gestión en España: Análisis de los artículos de la base de datos ISOC (2000-2009). *Revista Española de Documentación Científica*, vol. 36(2), 7.
- Bedia, A. M. S.; Fernández, M. C. L.; Pérez, M. P. (2013). Análisis de la relación entre flexibilidad en operaciones y performance empresarial mediante técnicas bibliométricas. *Revista Española de Documentación Científica*, vol. 36(4), 32.
- Berling, C. (2000). Continuous improvement as seen from groups and 'improvement agents'. *Total Quality Management*, vol. 11(4-6), S484-S489. <http://dx.doi.org/10.1080/09544120050007797>
- Cooney, R.; Sohal, A. (2004). Teamwork and total quality management: A durable partnership. *Total Quality Management & Business Excellence*, vol. 15(1), 1131-1142. <http://dx.doi.org/10.1080/1478336042000255442>
- Correa i Domènech, M.; García Blandón, J.; Úbeda Sales, R. (2009). Análisis de la producción científica en economía financiera: 1995-2006. *Revista Española de Documentación Científica*, vol. 32(2), 83-104. <http://dx.doi.org/10.3989/redc.2009.2.671>
- Dahlgard-Park, S. M.; Chen, C.; Jang, J.; Dahlgard, J. J. (2013). Diagnosing and prognosticating the quality movement - a review on the 25 years quality literature (1987-2011). *Total Quality Management and Business Excellence*, vol. 24(1-2), 1-18. <http://dx.doi.org/10.1080/14783363.2012.756749>
- De Leede, J.; Kees Looise, J. (1999). Continuous improvement and the mini-company concept. *International Journal of Operations & Production Management*, vol. 19(11), 1188. <http://dx.doi.org/10.1108/01443579910291087>
- González Alcaide, G.; Castellano Gómez, M.; Valde-rama Zurián, J. C.; Benavent, R. A. (2008). Literatura científica de autores españoles sobre análisis de citas y factor de impacto en biomedicina (1981-2005). *Revista Española de Documentación Científica*, vol. 31(3), 344-365.
- Higueta-López, D. (2011). Del manejo de personal a la dirección de personas. *Revista Innovar Journal. Revista de Ciencias Administrativas y Sociales*, vol. 21(40), 67.
- Houy, C.; Fettke, P.; Loos, P. (2010). Empirical research in business process management. Analysis of an emerging field. *Business Process Management Journal*, vol. 16(4), 619-661. <http://dx.doi.org/10.1108/14637151011065946>
- Jun, M.; Cai, S.; Peterson, R. T. (2004). Obstacles to TQM implementation in Mexico's maquiladora industry. *Total Quality Management*, vol. 15(1), 59-72. <http://dx.doi.org/10.1080/1478336032000149108>

- Ramos Rodríguez, A. R.; Ruiz Navarro, J. (2008). Base intelectual de la investigación en creación de empresas: Un estudio biométrico. *Revista Europea de Dirección y Economía de la Empresa*, vol. 17(1), 13-38.
- Ramos, A. R. (2004). *Estructura intelectual de la investigación en creación de empresas: Un estudio bibliométrico de su literatura científica 1956-2003*. (Tesis Doctoral) Universidad de Cádiz.
- Rapp, C.; Eklund, J. (2007). Sustainable development of a suggestion system: Factors influencing improvement activities in a confectionary company. *Human Factors and Ergonomics in Manufacturing*, vol. 17(1), 79-94. <http://dx.doi.org/10.1002/hfm.20064>
- Richart-Ramón, A.; Martínez-Blasco, M.; García-Blandón, J. (2011). Análisis de la producción científica sobre gobierno corporativo a través de ISI Web of Science. *Revista Española de Documentación Científica*, vol. 34(1), 79-101. <http://dx.doi.org/10.3989/redc.2011.1.776>
- Sanchez, L.; Blanco, B. (2014). Three decades of continuous improvement. *Total Quality Management & Business Excellence*, vol. 25(9-10), 986-1001. <http://dx.doi.org/10.1080/14783363.2013.856547>
- Serrano Bedia, A. M.; López Fernández, C.; García Piqueres, G. (2007). Gestión de la calidad en servicios: Una revisión desde la perspectiva del management. *Cuadernos De Gestión*, vol. 7(1), 33-49.
- Singh, J.; Singh, H. (2009). Kaizen philosophy: A review of literature. *The IUP Journal of Operations Management*, vol. VIII(2), 51-72.
- Singh, J.; Singh, H. (2012). Continuous improvement approach: State-of-art review and future implications. *International Journal of Lean Six Sigma*, vol. 3(2), 88-111. <http://dx.doi.org/10.1108/20401461211243694>
- Suárez-Barraza, M. F. (2008). Encontrando al kaizen: Un análisis teórico de la mejora continua. *Pecunia: Revista de la Facultad de Ciencias Económicas y Empresariales*, vol. 7, 285-311. <http://dx.doi.org/10.18002/pec.v0i7.696>
- Suarez-Barraza, M. F.; Angel Miguel-Davila, J. (2011). Implementation of kaizen in Mexico: An exploratory study for a Japanese managerial approach in the Latin American context. *Innovar-Revista de Ciencias Administrativas y Sociales*, vol. 21(41), 19-37.
- Warwood, S. J.; Roberts, P. A. B. (2004). A survey of TQM success factors in the UK. *Total Quality Management*, vol. 15(8), 1109-1117. <http://dx.doi.org/10.1080/1478336042000255460>



ESTUDIOS / RESEARCH STUDIES

Propuesta de un Centro de Archivo del Feminismo tras el análisis de los fondos documentales del Movimiento Feminista

María Adelina Codina-Canet*, Rosa San Segundo Manuel*

*Departamento de Biblioteconomía y Documentación. Universidad Carlos III de Madrid.
Correo-e: mariaadelina.codina@uc3m.es ; rosa.sansegundo@uc3m.es

Recibido: 10-02-2014; 2ª versión: 17-04-2015; Aceptado: 20-04-2015.

Cómo citar este artículo/Citation: Codina-Canet, M. A. y San Segundo Manuel, R. (2016). Propuesta de un Centro de Archivo del Feminismo tras el análisis de los fondos documentales del Movimiento Feminista. *Revista Española de Documentación Científica*, 39(1): e117. doi: <http://dx.doi.org/10.3989/redc.2016.1.1245>

Resumen: Las fuentes documentales primarias son esenciales para la investigación. El presente estudio realiza una revisión de las fuentes documentales de archivo y estudia la situación del caudal documental generado a partir de la lucha del movimiento de mujeres desde la Transición democrática española. Revelada la vulnerabilidad y desprotección de los archivos del feminismo y de mujeres, analizando las causas. Existe riesgo de desaparición de muchos de estos fondos como consecuencia de la naturaleza efímera tanto de los documentos como de los colectivos estudiados. Se fundamenta la propuesta de creación de un Centro de Archivo del Feminismo para salvaguardar y proteger los documentos que son patrimonio documental y parte de la memoria feminista y de la historia de las mujeres.

Palabras clave: Archivos privados; patrimonio documental; fondo documental; conservación; preservación; organizaciones feministas; colectivos de mujeres; feminismo; género.

Proposal of a Feminist Archive Center after analysis of the documentary resources of the Feminist Movement

Abstract: Primary sources are essential for scientific research. This study reviews the documentary sources and analyses the state of the flow of documents generated from the struggle of the women's movement since the Spanish Democratic Transition. It reveals the vulnerability and the lack of archives on feminism and women, as they are outside the scope of protection afforded by archival systems. There is a risk that many of these papers could disappear because of the ephemeral nature of the groups studied. The analysis leads to a proposal for the establishment of a Feminism Archive Centre, to safeguard and protect documents that represent bibliographic heritage as well as the memory of the feminist movement and women's history.

Keywords: Private archives; documentary heritage; antique documentary; preservation; perpetuation; feminist organization; collectives of women; feminism; gender.

Copyright: © 2016 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0.

1. INTRODUCCIÓN

La pregunta se refiere a lo que está ocurriendo actualmente con los documentos de archivo, fuentes primarias para la investigación, generados por el movimiento feminista a partir de la década de los años 70. Este trabajo analiza el estado de la cuestión, planteando soluciones. Es una toma de conciencia en la puesta en valor del Patrimonio Documental Feminista, como fuente documental necesaria para posteriores investigaciones de movimientos feministas recientes en España y la reconstrucción histórica del asociacionismo de mujeres en ese período.

El valor histórico de los documentos de archivo es de gran importancia, en parte amparado y protegido por la Ley de Patrimonio Histórico Español y por la Estrategia y Programa del Sector de Comunicación e Información de UNESCO. A partir de ello se repara en la necesidad de salvaguardar los archivos personales y de grupos organizados de mujeres y feministas desde su surgimiento, teniendo en consideración que en bastantes casos se trata de colectivos o grupos de vida efímera. Además de ser corrientes las escisiones y divisiones generando la dispersión inicial en diversos grupúsculos. Surgen algunas preguntas: ¿Qué cantidad de archivos y fondos documentales se están tratando? ¿Dónde se localizan?, ¿Cómo se conservan y gestionan los fondos documentales? ¿Cómo influyen en la investigación y en la historia?

En 1996, con motivo de la conmemoración de la mayoría de edad de la Constitución de 1978, activistas de colectivos y grupos de las luchas por los derechos de las mujeres en la Transición española, pusieron de manifiesto las notorias ausencias en las obras de historia de lo referente al feminismo español. A partir de esta exclusión se decide actuar, y se deja constancia de la importancia que, para una sociedad como la española de esos años de transformación, tuvo la participación de las mujeres en grupos, asociaciones y partidos, resaltando el valor de los testimonios de su protagonismo político.

La célebre jurista María Telo Núñez (1915-2014), que tanto en la etapa franquista como en los primeros años de la Transición lideró la lucha contra la discriminación jurídica de las mujeres y, a nivel internacional es considerada un referente del feminismo español de la época (Pazos, 2014), era muy consciente de la omisión de las mujeres en la historia reciente. Telo interesada por el libro "Historia de las mujeres en España" de las investigadoras Elisa Garrido, Pilar Folguera Crespo, Margarita Ortega y Cristina Segura, afirmó: *su lectura me ha llevado a constatar que donde se trata del fi-*

nal del franquismo y de la democracia y cambio social, hay una gran laguna que nos afecta a las mujeres juristas que tanto luchamos por erradicar del Derecho de Familia (Código Civil y Código de Comercio) el trato tan discriminatorio que la mujer sufría (Telo Núñez, 1998).

A pesar de estas omisiones fue determinante la influencia que ejercieron los diferentes grupos feministas y los sectores profesionales más progresistas en los cambios legislativos que afectaban al comportamiento sexual de las mujeres, pues eran objeto de un tratamiento punitivo exclusivamente para las mismas (Folguera Crespo, 1997).

En otros estudios de género relacionados se ponen de manifiesto las mismas omisiones (Anderson, 1991), ya que en el periodo de la Transición existió una lucha organizada, en muchas ocasiones con campañas unitarias por los cambios legislativos, pero que al narrarse se han invisibilizado. Sin embargo, han cobrado mayor visibilidad aquellas acciones realizadas por grupos feministas adscritos a los partidos políticos (Cuesta Bustillo, 2003), y fueron especialmente relevantes los grupos feministas organizados en relación a las reformas políticas y legislativas de la Transición, a pesar de que su legado nos haya llegado más de las organizaciones feministas incluidas en los partidos políticos.

Es muy importante evitar que estas prácticas exclusivistas no se tornen en una imposición generada por las siguientes investigaciones. Para la construcción de una historia más veraz e inclusiva es necesario poner al servicio de la investigación los documentos generados por el movimiento feminista organizado. Hay que comprender que los documentos son la huella que dejamos en la memoria para que ésta perdure y nos supere en el tiempo.

Los fondos documentales que se generan en el desarrollo de las actividades de lucha por los derechos de las mujeres, sean archivos personales o de organizaciones feministas, como material documental de archivo ha de ser conservado porque en ellos se encuentra la huella del pasado, las identidades presentes y el futuro de la historia. En este sentido, destaca Alberch Fugueras y otros, (2001): *los archivos colaboran en la asunción de los valores del patrimonio público, la memoria, la identidad y el conocimiento.*

Se incluyen en este frente documental para la investigación con perspectiva de género las distintas tipologías documentales tradicionales de un fondo de archivo, trátense de manuscritos, correspondencia, materiales gráficos, agendas, diarios y cualquier documento no editado, que constituyen las fuentes primarias de la historia de las mujeres. Por tanto, si son tratados desde la epistemología

archivística, van a contribuir en un futuro a escribir la aportación intelectual a la historia de la lucha feminista y a la construcción de la memoria colectiva e histórica de las mujeres.

Resulta preciso preguntarse el estado de vulnerabilidad en el que se encuentra el caudal documental inédito feminista, qué centros públicos o privados se encargan de su custodia, así como su accesibilidad a la información como elemento esencial y fuente documental para la investigación científica del movimiento feminista y la historia de las mujeres.

La archivera Cortés Alonso (1986), analiza en las novelas de Emilia Pardo Bazán, cómo se aborda el tema de los documentos y los archivos en sus relatos novelados. En los mismos aparecen los fines, la instalación, los fondos documentales, la conservación y la organización del archivo en su tiempo. Cortés caricaturiza, en su caso, las guías e inventarios elementales descritos por Bazán. Bien distinto, nos explicita, de *los instrumentos de información normalizados, claros y fáciles de manejar, que honran a quienes los hace y beneficia a quien los consulta*. Cortés extrae en su artículo numerosas citas de la obra, para analizar y recargar los conceptos archivísticos actuales y su importancia, por su aplicación práctica, y por el rigor científico. Concluirá sobre lo que narra la ilustre escritora: *Este tema de archivo es un pequeño accidente en la trama de la novela -Los pazos de Ulloa-. Sin duda, pero es una muestra de las lecciones que iba introduciendo, la autora en sus obras, para reformar defectos y mejorar el estudio cultural y material de sus lectores ...; en el paso de la ficción novelada a la realidad cotidiana, del ejemplo literario y didáctico al caso habitual del servicio archivístico en nuestros centros. Casi cien años transcurridos y la lección es práctica y actual. Vale la pena anotarlo*. Los archivos son imprescindibles para la anotación de nuestras memorias.

En todo caso, para salvaguardar el patrimonio documental del feminismo disperso en la actualidad, habría que tomar hoy las medidas necesarias para evitar en unos años el tipo de preguntas planteadas por la investigadora Caballé Masforroll (2006), en relación a los estudios biográficos: "¿Cómo hubieran sido las cosas de conservar las cartas escritas por Rosalía de Castro y destruidas a su muerte por su esposo Manuel Murguía? ¿Dispondríamos de un conocimiento más directo y personal de la escritora y jurista Concepción Arenal si, por el pánico que sentía a las murmuraciones, no hubiera quemado con determinación metódica sus documentos personales?".

Un caso fortuito y cercano en el tiempo es el archivo de la Residencia de Señoritas de Madrid (1915-1936), tras encontrarse hace tan sólo unos años de manera casual en un estado de abandono tal que si no se hubiese intervenido inmediatamente hubiese sido destruido (Capel Martínez, 2009). Hubiera sido una gran pérdida ya que cualquier investigación científica sobre la historia de una institución debe reunir diversas fuentes de información, incluyendo la documentación de la propia institución (Codina-Canet, 2013) *Las fuentes de archivo, ordenadas cronológicamente aportan consistencia a la realidad histórica y de la institución objeto de estudio. Permitiendo describir en su conjunto lo trascendente en la historia*. Sería conveniente que las asociaciones de mujeres no repitiesen errores y tomaran conciencia de la necesidad de conservar su documentación, material imprescindible para estudiar la historia reciente de la lucha de las mujeres.

2. OBJETIVOS

El objetivo principal de este trabajo es conocer la situación de los fondos documentales que serán constitutivos de un archivo feminista originados a partir de la Transición política en España. El ámbito geográfico y temático abarca federaciones de asociaciones y asociaciones de ámbito estatal, que se centralizan en Madrid. Se trata de un estudio teórico que establece los fundamentos para una propuesta de creación de un centro de archivo temático, aunque esta no va acompañada de desarrollo alguno que permita su viabilidad. Para las situaciones de vulnerabilidad de los fondos se proponen soluciones de viabilidad para la custodia de los mismos. Para estos fines se establecen los siguientes objetivos de investigación como:

- Reunir fuentes documentales de primera mano, para la investigación interesada en el movimiento feminista.
- Identificar los productores representativos de fondos documentales feministas en relación a su posible integración en Archivos estatales o privados.
- Plantear el asociacionismo como estrategia para dar solución a los problemas, que garantice la protección del caudal documental del feminismo, mediante el establecimiento de redes entre instituciones del feminismo académico y centros de investigación y documentación.
- Proponer un sistema de archivo del feminismo y el centro de archivo temático como solución para salvaguardar los documentos de archivo de naturaleza feminista.

Finalmente, mostrar el justo valor que deben tomar las organizaciones feministas de sus propios documentos, en la medida que han de constituir el patrimonio documental del feminismo español. E igualmente, facilitar la toma de conciencia de la situación por parte de quienes investigan el feminismo en España. En definitiva, sirva esta investigación para fundamentar actuaciones de intervención o colaboración de las Administraciones Públicas responsables y garantes en la protección del Patrimonio Documental español.

3. MATERIAL Y MÉTODOS

La metodología pasa por reconocer un volumen potencial e incierto de agentes productores de documentación y articularlos para el centro de archivo temático futuro según el concepto de principio de procedencia y clasificación de documentos (Herrero Montero y Díaz Rodríguez, 2011). Agentes productores representados por grupos, colectivos, asociaciones del movimiento feminista y otros. Existe gran desconocimiento de las cantidades de documentos de archivo que se conservan, de la organización archivística de los mismos, y por tanto, del servicio público que pueden estar prestando a la comunidad investigadora como fuente de documentación.

En primer lugar, se expone la delimitación temporal y terminológica utilizada, y a continuación la estrategia seguida en la investigación. Finalmente, las fuentes tanto bibliográficas como documentales existentes.

El elemento de análisis se ha denominado *archivo feminista* constituido por los documentos de archivo producidos por agentes en sus actividades de lucha por los derechos de las mujeres. Se precisa una limitación temporal y terminológica para así centrar el objeto de estudio. El análisis se restringe a los documentos producidos por los colectivos, grupos y asociaciones feministas desde el inicio del proceso de reinstauración de la democracia española, a mediados de los años 70, y comprende hasta las dos décadas siguientes.

En relación a los vocablos empleados, se designa "archivo", en minúsculas, al contenido documental y "Archivo" como edificio custodio del contenido.

Con referencia al concepto de *archivo feminista*, se trata de un concepto ausente en la literatura científica española en el campo de la Archivística. La unión de ambos vocablos archivo y feminista, una concatenación necesaria de términos, se encuadra por el momento, en el grupo de archivos privados. Es el constituido por fondos documentales de archivo que produce y recibe una persona,

una entidad o colectivo en el ejercicio del desarrollo de funciones y actividades feministas.

De las muchas acepciones de archivo que se encuentran en manuales y tratados de archivística servirá la establecida en el Sistema Español de Archivos (R.D. 1708/2011): *archivo incluye tanto el conjunto orgánico de documentos, o la reunión de varios de ellos, producidos o reunidos por las personas físicas o jurídicas, públicas o privadas. Como entidades, que de acuerdo con las normas internacionales de descripción archivística, comprenden instituciones, personas y familias, que reúnen, conservan, organizan, describen y difunden los conjuntos orgánicos y las colecciones de documentos.*

Finalmente el vocablo feminismo se va a emplear desde un análisis abierto, no excluyente, no uniforme y no lineal de definición del feminismo como movimiento plural de múltiples itinerarios y estrategias de emancipación femenina (Nash, 1994). El término feminismo en tanto que movimiento organizado que parte de la existencia histórica de discriminación de la mujer, y que implica un proceso de transformación social y cultural. Y activista será aquella persona o colectivo implicado en este movimiento social.

La metodología seguida se articula en distintos elementos de análisis, relacionándose y vinculándose entre sí, ahondando en el perfil del problema, y a través de la discusión se obtendrán unas conclusiones que permitan definir soluciones y mejoras del problema.

El primer elemento de análisis son los agentes productores o instituciones del movimiento feminista, así como su identificación y análisis en función del carácter organizativo, lo que va a abarcar a colectivos feministas profesionalizados dedicados a la ayuda y a la asistencia social, jurídica, psicológica, mediación familiar y otros. También abarca a colectivos y grupos feministas autónomos, inclusive organizaciones de encuentros de mujeres. Colectivos feministas de doble militancia, formados por grupos integrados en partidos políticos o grupos feministas adscritos a algún sindicato, y finalmente, agentes de carácter individual, como las mujeres feministas.

El segundo elemento de análisis se relaciona con los fondos documentales en relación a la integración en centros de Archivo. En relación a la valoración de los fondos localizados, esta investigación no procede a la valoración de los documentos, ya que se trata de una propuesta teórica y no contempla una descripción documental. Con respeto al principio de la trascendencia del documento, se considera, a priori, que un fondo documental cuyo

productor es un colectivo feminista, es condición suficiente para dotarle de valor patrimonial único y de especial protección por su interés para la investigación y la historia (Vázquez Murillo, 2006), con independencia del valor documental administrativo, jurídico, fiscal, informativo o único e histórico que tuviese, cuestión que quedaría resuelta tras la aplicación de técnicas archivísticas.

3.1. Fuentes

Para abordar la investigación de los agentes productores de fondos documentales, se va a indagar en censos, guías y revistas de la temática, con un estudio bibliográfico que proporcione unas bases sólidas en la investigación mediante una revisión bibliográfica y de fuentes de archivo vinculadas al movimiento feminista de la Transición. Se va a proceder a una selección de investigaciones de esta temática que incluyan contenidos referentes a los grupos, organizaciones y colectivos del movimiento feminista y fuentes documentales de archivo.

Se trata de trazar un plan de localización de estos fondos de archivo y su consulta en centros de Archivos públicos y privados. Las fuentes documentales agrupadas en tres categorías: fuentes generales, fuentes de centros de Archivo y documentación y fuentes bibliográficas, así como bases de datos de archivos digitales sobre mujeres, feminismo y otros temas afines.

En relación a los centros de Archivo que puedan poseer depósitos de grupos y asociaciones feministas, se van a excluir el Archivo General de la Administración que custodia fondos que han sido producidos en su mayor parte por la actividad de los organismos públicos en el ámbito de la Administración Central, y los centros de Archivo de Diputaciones, porque estos custodian documentos cuya procedencia es la Administración Local. Ciertamente en los citados centros se podrían encontrar expedientes administrativos sobre la temática, pero se van a excluir ya que no contemplan fondos de archivo producidos por Organizaciones no Gubernamentales de esta época reciente.

Los primeros resultados obtenidos principalmente de la recuperación de información de las bases de datos, con acceso abierto y, en otros casos, presencialmente en catálogos e inventarios, son muy insuficientes. Se ha hecho necesario introducir fuentes de información con el sistema seguido de sondeo. Método que abarca peticiones de información mediante formularios y entrevistas, con testimonios orales para poder ampliar los datos del estudio de manera significativa en la localización de archivos personales. A este respecto las organizaciones integradas en redes académicas relacio-

nadas con la documentación del feminismo y los centros de la Red de Centros de Documentación y Bibliotecas de Mujeres, con relación al feminismo académico, las entidades afiliadas a la Asociación Universitaria de Estudios de Mujeres, en adelante AUDEM, la documentación e información al respecto era muy escasa, por lo que mayoritariamente la tenencia de fondos de archivo surge de colectivos y asociaciones del movimiento feminista.

Ante la información ocasionalmente parcial e insuficiente en la localización de fondos documentales, son de interés las investigaciones para la estimación de las fuentes documentales de publicaciones cuyo objeto de estudio es el feminismo en la Transición democrática (Anderson, 1991; Folguera Crespo, 1997, Asociación "Mujeres en la Transición Española", 1999; Cuesta Bustillo, 2003; Agustín Puerta, 2003; Nash, 2007; Martínez Ten y otros, 2009 y Rubio, 2012). Destaca Agustín Puerta (2003), en tanto que aborda en su investigación las fuentes primarias de archivo.

Estas fuentes de información fundamentan la necesidad de creación de un centro de fondos del feminismo, para lo que sería necesaria una recopilación de fuentes de archivo exhaustiva, el propio centro habrá de localizar y recopilar los archivos dispersos del feminismo.

Sobre fuentes documentales generales destacan Bases de datos Censo-Guía de Archivos de España e Iberoamérica, Portal de Archivos Españoles, Guía de fuentes documentales de archivo y base de datos del Archivo de la Biblioteca Nacional.

Sobre fuentes de centros de archivos de titularidad estatal destacan: el Archivo Histórico Nacional y el Centro Documental de la Memoria Histórica.

Sobre archivos históricos provinciales o de Comunidades Autónomas destaca: el Archivo Histórico Nacional de Catalunya.

Con relación a los Archivos de universidades españolas destacan el Archivo General de la Universitat d'Alacant, el Arxiu Històric de la Universitat de Barcelona, el Archivo Histórico de la Universidad de Oviedo, el Archivo General de la Universidad Complutense, el Archivo Universidad Autónoma de Madrid y el Archivo General de la Universidad Carlos III.

Entre los centros universitarios de Estudios de Mujeres son relevantes el Centres d'Estudis sobre la Dona, Universitat de Barcelona, el Institut Universitari d'Estudis de la Dona de la Universitat de València, el Grupo de Investigación de la Universidad de Oviedo, Deméter, el Centre Dona i Literatura. Gènere, sexualitats, crítica de la cultura, de la Universitat de Barcelona, el Instituto

Universitario de Investigación Feminista y de Género Purificación Escribano, Universitat Jaume I, el Instituto de Investigaciones Feministas, Universidad Complutense de Madrid y el Instituto de Estudios de Género de la Universidad Carlos III.

En relación a los centros de documentación y bibliotecas de mujeres destacan el Emakumeen Dokumentazio Zentrua (Bilbao), el Centre de Documentació del Institut Català de les Dones, la Biblioteca de Mujeres (Madrid), el Centre de Documentació Ca la Dona (Barcelona), la Biblioteca Casa de la Dona (València) y el Centro de Documentación e Recursos Feministas de Vigo.

Sobre archivos históricos de fundaciones, partidos políticos y sindicatos obreros destacan el Archivo Histórico de la Fundación 10 de Marzo, de Santiago de Compostela, el Archivo Histórico de Comisiones Obreras de Andalucía, el Archivo Histórico de Comisiones Obreras de Cataluña, el Archivo Fundación Pablo Iglesias, el Archivo Histórico del Partido Comunista y el Archivo Documental de la Liga Comunista Revolucionaria

No existen fuentes bibliográficas específicas por lo que se hace necesaria una selección bibliográfica del feminismo de la transición. El primer problema resultante de la valoración del patrimonio feminista documental es la ausencia de publicaciones de guías o fuentes de información específicas de los fondos documentales de archivo del feminismo en España. Los resultados presentan el análisis de los grupos y organizaciones como agentes productores de documentación, estableciendo una clasificación operativa para el estudio de los mismos.

4. GRUPOS Y ORGANIZACIONES FEMINISTAS AGENTES PRODUCTORES DE DOCUMENTACIÓN

El asociacionismo de las mujeres en grupos y organizaciones feministas durante el período de transición y primeros años de la democracia (1975-1985) se caracteriza por ser muy prolífico en su lucha organizada por la emancipación de las mujeres. Una actividad incluso febril por la coyuntura política provocada ante la perspectiva del nuevo texto constitucional de 1978. La mayoría de grupos, en 1976, se constituían en la unidad de acción reivindicativa de la Plataforma Unitaria de Organizaciones y Grupos de Mujeres de Madrid, que llegó a integrar a 55 grupos o más. A nivel estatal sería la Coordinadora Estatal de Organizaciones Feministas la que aunó la acción actuando como un agente colectivo para conseguir cambios sociales.

En la primera guía de 1987 sobre instituciones y asociaciones de mujeres en España (Guía de asociaciones, 1996), se registran cerca de 1000 aso-

ciaciones y grupos. El cómputo incluye a las asociaciones de mujeres, y grupos de organizaciones de diferente naturaleza que trabajan con mujeres de manera prioritaria, pero que, sin embargo, su actividad es otra, como fueran asociaciones de vecinos, ateneos y otros, que incluyen en su trabajo a los colectivos femeninos. Los datos se elaboraron en función de las peticiones de subvención dirigidas al Instituto de la Mujer creado en 1983. Por tanto, se excluyen del cómputo los grupos organizados más autónomos e independientes (Guía de asociaciones, 1987). En esta relación, al desagregar los datos por Comunidades Autónomas, resulta ser la más numerosa Andalucía, con 126 asociaciones, seguida de Madrid con 108, a Castilla y León le corresponden 85 y a Catalunya 83 entidades.

Los Colectivos Feministas que van surgiendo en el Estado español, citados en la publicación seriada *Vindicación Feminista*, referente más relevante del feminismo durante la Transición (Los Colectivos Feministas, 1976), dan cuenta del activismo social establecido en el momento. Las raíces del feminismo organizado se encuentran en el proceso de lucha contra la dictadura, raíces paralelas a los partidos políticos clandestinos en la oposición al régimen, y posterior proceso de la Transición política. Existieron unos 180 grupos considerados más activos, asociaciones, coordinadoras y colectivos feministas (Augustín Puerta, 2003). En la Guía de asociaciones de mujeres de la Comunidad de Madrid, en la que consta que ésta agrupa en ese momento a 57 asociaciones pertenecientes al Consejo de la Mujer y 262 asociaciones no pertenecientes al citado Consejo, computando en total más de 300 grupos y organizaciones (Guía de asociaciones, 1996). En la asociación "Mujeres en la Transición Democrática" también se recopiló la participación de las mujeres en grupos de asociaciones y partidos políticos. Las feministas activas fueron mujeres que conformaron nuestra historia. Algunos de los colectivos ubicados en Madrid fueron muy representativos en la Transición española como el Seminario de Estudios Sociológicos sobre la Mujer, Movimiento Democrático de Mujeres, Asociación Española de Mujeres Universitarias, Asociación de Amas de Casa y Hogar, Federación Asociaciones de Amas de Casa, Fundación Centro de Investigaciones y Formación Feminista, Asociación para la Promoción y Evolución Cultural, Asociación Democrática de la Mujer, Federación de Organizaciones Feministas, Asociación Española de Mujeres Juristas, Asociación Española de Mujeres Separadas, Seminario Colectivo Feminista y Colectivo Feministas de Madrid, Colectivo Jurídico Feminista, Frente de Liberación de la Mujer, Plataforma de Organizaciones Feministas y Grupos de Mujeres de Madrid, Asociación Universitaria

para los problemas de la Mujer, Mujeres de Acción Católica, Asociación de Madres Solteras para la Igualdad, Asociación Castellana de Planificación Familiar, Federación Nacional de Asociaciones de Planificación Familiar, Asociación Española de Planificación Familiar, Asociación Española de Mujeres Empresarias, Club Vindicación Feminista, Asociación Mujeres por Europa, Asociación Democrática Conservadora y Federación Nacional de Mujeres para la Democracia.

El Acta de la reunión de la primera Comisión consultiva para el seguimiento del Plan para la Igualdad de Oportunidades de las Mujeres (1988-1990), celebrada en 1988, cita a 100 colectivos de todos los territorios del Estado, en tanto que participantes de esa Comisión (Acta de la reunión, 1989). Hay constancia en relación a las ayudas y subvenciones del Gobierno de la Comunidad de Madrid a asociaciones de mujeres y ayuntamientos en 1990 donde se citan a 233 grupos de asociaciones, siendo beneficiarias de la misma, 132 asociaciones, de ellas 89 se ubican en la capital (Subvenciones a asociaciones, 1990). El Informe, recientemente publicado, de recuperación de la memoria de grupos, colectivos y organizaciones feministas en Galicia, recoge 226 organizaciones (Pérez Pais y Mariño Costales, 2013).

5. PROPUESTA CLASIFICATORIA PARA EL ESTUDIO DE LA DOCUMENTACIÓN DEL FEMINISMO

La protección de los documentos de archivo requiere diferenciar las distintas asociaciones según sea la naturaleza organizativa de la entidad o grupo analizado, en función de la naturaleza organizativa de cada una por el tipo de actividades de reivindicación y lucha feminista. La protección, clasificación y perdurabilidad de la masa documental, en función de la entidad o grupo analizado. Esta diferenciación está basada en la casuística utilizada por Agustín Puerta (2003) en la investigación del origen y evolución del feminismo en la Transición, donde se diferencia entre entidades autónomas, profesionalizadas, doble militancia y entidades unipersonales.

- Entidades de organizaciones feministas autónomas, caracterizadas por grupos de mujeres con intereses y objetivos estratégicos, con una conciencia específicamente feminista, en un abanico de posibilidades que van desde los grupos independientes, a los organizados en barrios, grupos de autoconciencia y grupos reivindicativos de cambios de legislación en materias de divorcio y separación, salud, familia y violencia de género. Los fondos localizados se encuentran en los centros de documenta-

ción y bibliotecas especializados en mujeres y feminismo cuyo referente de creación son el movimiento feminista, y también en algunos Archivos universitarios.

- Colectivos feministas profesionalizados, grupos constituidos por asociaciones dedicadas a la ayuda legal, psicológica, y a la asistencia social, y las entidades para la formación e investigación feminista.
- Colectivos feministas de doble militancia. Grupos feministas integrados en partidos políticos y grupos feministas adscritos a sindicatos obreros.

El problema de recuperación de fondos se agudiza cuando se trata de colectivos feministas autónomos y, en mayor medida, cuando su naturaleza fuera efímera, ya que muchos de ellos ya están desaparecidos. Sin embargo, al tratarse de actividades en la historia reciente, numerosas feministas que pertenecieron a ellos pueden haber guardado parte de la documentación generada. Existieron grupos muy activos, y aquellos que contaron con más recursos, editaron publicaciones periódicas propias para la difusión de sus actividades e idearios.

En relación al registro documental de las actividades de los grupos, ha habido mayor conservación documental cuando eran receptores de subvenciones gubernamentales, ya que ello generaba documentación administrativa justificativa de las ayudas tales como memorias de actividades, contabilidad y actas de reuniones, donde constaba el activismo y se facilitaba la descripción de su identidad. Las entidades profesionalizadas tuvieron mayor posibilidad de perdurar, se trata de grupos generalmente dotados de personal administrativo que se ocupaba de archivar los documentos que se generaban en la organización y en el trato con las usuarias.

En cuanto a los grupos feministas integrados en organizaciones como partidos políticos, sindicatos, universidades, instituciones gubernamentales, en las que existe un sistema organizativo y, por consiguiente, documental archivístico y una mayor infraestructura, los fondos documentales de estos grupos feministas, que han trabajado desde la denominada doble militancia, tiene menor grado de amenaza de desaparecer.

5.1. Fondos documentales del feminismo en centros de archivo

Los focos del feminismo organizado se dan en todo el territorio del Estado español, con notables diferencias entre Comunidades Autónomas y, sin embargo, las asociaciones de ámbito estatal se

centralizan en Madrid. Es notoria, la dispersión y deslocalización de fondos del feminismo, debido a la inexistencia de una entidad institucionalizada abierta e integradora para albergar y proteger el patrimonio documental de esta temática.

Existe un depósito mínimo de documentos de la temática en centros integrados en el Sistema Español de Archivos: el Archivo Histórico Nacional y el Centro Documental de la Memoria Histórica. Y en mayor medida en otros centros de Archivo externos al mismo: el Archivo Histórico Nacional de Catalunya, los centros de Archivo universitarios, los centros de documentación y bibliotecas especializados en mujeres y feminismo y, finalmente, los centros de Archivo histórico de partidos políticos, sindicatos y fundaciones políticas.

5.1.1. Archivos de la Administración

Los centros de Archivo españoles que conforman el Sistema Nacional Archivístico son una red de centros que conserva el Patrimonio Histórico Documental Español. Los fondos documentales localizados en estos centros son escasos. El Fondo del Centro de Investigación y Formación Feminista CIFFE se conserva depositado en el Centro de la Memoria Histórica. Próximamente se incorporará el archivo personal de María Telo y de la Asociación de Mujeres Juristas al Archivo Histórico Nacional (Telo Núñez, 2009).

En el Archivo Histórico Nacional de Catalunya están depositados los archivos personales de Lidia Falcón, Montserrat Roig y María Dolors Calvet Puig.

5.1.2. Archivos en las universidades españolas

Actualmente hay depositados fondos de activistas del feminismo en el Archivo Histórico de la Universidad de Oviedo y en el Archivo General de la Universitat d'Alacant. El fondo de Asociación de Feministas Asturianas (1960-1990), donado por Lourdes Pérez González. En el caso de la Universidad de Oviedo, la tesis de Carmen Suárez, que estudia el feminismo en Asturias, rescata un conjunto de ocho archivos personales que han incrementado el archivo primigenio de Asociación Feminista Asturiana (AFA). Refiriéndose a archivos personales originados en la Transición española de las feministas Begoña Sánchez, M^a José del Río, Paloma Uría, Oliva Blanco, Consuelo Cambor, María Martínez Rodríguez y M^a Luisa Rodríguez Fernández (Suárez Suárez, 2005). En la Casa de Encuentros de las Mujeres de Gijón permanecen dos archivos aún inexplorados: el Fondo de Paz Fernández Felgueroso y el de Dulce Gallego Canteli (Suárez Suárez, 2012). El

Fondo Archivo de la Democracia de la Universitat d'Alacant contiene documentos donados por Inmaculada Fernández Arrillaga y M^a Teresa Molares Mora.

Desde 2005 se trabaja, en distintas universidades, en proyectos de archivos orales, en este sentido destaca el Archivo de Fuentes Orales para la Historia Social de Asturias (AFOHSA) de la Universidad de Oviedo que comprende un número importante de entrevistas, relatos de vida, testimonios de represión y de violencia política en Asturias relativos a las huelgas de 1962 y las "culturas del trabajo" y otros temas afines. En el AFOHSA constan los testimonios orales de Anita Sirgo Suárez y de Ana García Carpintero en la serie Disertaciones y de Aida Fuentes Concheso, de Francisco Fernández Corte, de Ángeles Pollo Joglar y de Laura González Álvarez en la serie Historias de Vida.

En los centros de investigación feminista de las universidades se localizan algunas donaciones, como el fondo documental de Anna Bosch i Pareros donado en 1991 al Centre d'Investigació d'Història de la Dona de Barcelona (Elvira i Silleras, 2013), actualmente conocido como DUODA. El fondo personal de Presen Sáez de Descatllar donado al Institut Interuniversitari d'Estudis de la Dona de la Universitat de València (IUED).

5.1.3. Centros de documentación y bibliotecas de mujeres

En cuanto a los centros de documentación y bibliotecas de mujeres son considerados unidades especializadas de información de mujeres y feminismo. Ante los primeros cambios después del franquismo, y la demanda social van a emerger los primeros centros. Para las investigadoras Muñoz-Muñoz y Argente Jiménez (2010): "...su aparición tiene lugar a finales de la década de los años 70, al amparo de la nueva situación política y fruto del interés de asociaciones e instituciones". Estos centros van a ser receptores de una demanda, que se iniciará con la formación de los primeros fondos documentales sobre mujeres en organizaciones, colectivos y asociaciones del movimiento feminista, los centros de investigación universitaria y los organismos gubernamentales promotores de políticas públicas de igualdad entre mujeres y hombres en la década que se inicia en 1980. Se trata de tres naturalezas distintas de centros: los pertenecientes a colectivos y asociaciones del movimiento feminista, los centros de investigación universitaria y los organismos gubernamentales: Instituto de la Mujer y Centros de Información de Ayuntamientos.

En la guía de asociaciones de mujeres y de centros de interés (Guía de asociaciones, 1989), se citan 11 centros: Centro de Documentación del Instituto Andaluz de la Mujer (Málaga), Centro de Documentación de la Mujer (Zaragoza), Centre de Documentació de la Dona (Barcelona), Centre de Estudios de la Mujer (Barcelona), Departamento Interdisciplinar de Estudios de la Mujer IPES (Pamplona), Centro de Documentación Instituto Vasco de la Mujer (Álava) y Centro de Documentación del Institut Valencià de la Dona (Valencia), Biblioteca de Mujeres (Madrid), Centro de Documentación del Instituto de la Mujer (Madrid), Centro de Investigación y Formación Feminista (Madrid), Centro Feminista de Estudios y Documentación (Madrid), los cuales fueron creados con el esfuerzo de los grupos de mujeres preocupadas por la recopilación de fondos bibliográficos y documentales. De ellos, cuatro se ubicaban en Madrid. Actualmente están activos el Centro de Documentación del Instituto de la Mujer y la Biblioteca de Mujeres (Madrid) que surge en 1985, como biblioteca de conservación del movimiento feminista (Marisa Mediavilla, entrevista personal, 5 de marzo de 2014).

Desde 1993, estos centros y otros creados posteriormente, cooperan a través de la Red de Centros de Documentación y Bibliotecas de Mujeres. La mayoría de centros de la Red carecen de fondos de archivo del feminismo, con la excepción de los centros que surgen en la base de colectivos y asociaciones del movimiento feminista.

En 1998 se creó la Asociación de Centros de Documentación y las Bibliotecas de Mujeres "María Moliner", de ámbito estatal constituida por aquellas bibliotecas y centros de documentación que han surgido del Movimiento Feminista (Fraile, 2000).

Se localizan fondos documentales de archivo en los siguientes centros de documentación y bibliotecas de mujeres: Ca la Dona (Barcelona), Casa de la Dona (Valencia), el Centro de Documentación de Mujeres Maite Albíz (Bilbao), el Centro de Documentación y Recursos Feministas de Vigo y la Biblioteca de Mujeres (Madrid). Todos estos fondos de archivo localizados en centros de documentación y bibliotecas de mujeres son consecuencia directa de compartir el espacio físico con colectivos autónomos, siendo centros de referencia para recibir donaciones individuales o colectivas. Algunos de grupos o asociaciones desaparecieron y con ellas también han desaparecido sus documentos por no existir un lugar apropiado al que donarlos. Los fondos documentales de archivo de estos centros no han sido tratados con metodología archivística, desconociéndose la existencia de cuadro de clasificación, inventarios o catálogos documentales.

La Biblioteca de Mujeres (Madrid), además del fondo documental propio contiene en depósito documentos de gran interés de grupos, la mayoría de ellos ya desaparecidos: Unión para la Liberación de Mujeres, Movimiento Democrático de Mujeres, Asociación Democrática de Mujeres, Hypatia, Walkirias, Red Europea de Mujeres, Federación de Asociaciones de Mujeres Flora Tristán, Colectivo de Lesbianas, Asociación de Amas de Casa de Tetuán, Centro de Investigación y Formación Feminista (CIFFE), entre otros. Estos fondos documentales de archivo son citados y referenciados en la investigación de Augustín Puerta (2003).

El Centro de Documentación Ca la Dona, además de archivo gráfico y administrativo propio, contiene los fondos personales de Gretel Amman Martínez, Encarna Sanahuja, Carme Alemany y Victòria Sau, Grupo L.I.C.I.T. (Línia d'Investigació i Cooperació amb Immigrants i Treballadores Sexuals), el fondo documental del Proyecto Vaca y Associació de Creadores Escèniques. Todos estos archivos presentan una colección de documentos que tienen en común la actividad propia de la institución (sea grupo o persona) preparada profesionalmente como testimonio para la memoria de la institución (Mastropierro, 2006).

Por lo que respecta a los centros de documentación especializados y bibliotecas de mujeres que cuentan con depósito de fondos de archivo, sin embargo, carecen de estructura de archivo. Algunos, que por su proximidad al movimiento feminista y disponer de mayor infraestructura, tienen depositados varios de sus fondos en ubicaciones localizadas, se esfuerzan voluntariamente en poner a disposición de la investigación sus fondos, reconociendo la insuficiencia de medios para ofrecer un servicio de acceso adecuado, ya que carecen de proyecto archivístico tanto físico como digital viable por falta de financiación, además de carecer de una organización física estructurada de los fondos y de instrumentos de descripción archivística.

5.1.4. Centros de archivos privados de Partidos Políticos, Sindicatos y otras Fundaciones

A finales de la década de los años 70, comienzan a proyectarse los Centros de Archivos privados de Partidos Políticos, Sindicatos y otras Fundaciones políticas históricas que sobrevivieron a los tiempos de la clandestinidad. Estos centros custodian fondos de grupos feministas de doble militancia. Al tener mayores infraestructuras, están dotados de estructura organizativa archivística. Estos archivos han evolucionado de manera proactiva, y actualmente se agrupan en redes, como la Red de Archivos Históricos de Comisiones Obreras, los Archivos de la Unión General de Trabajadores, y otros.

Los fondos personales de feministas de doble militancia están siendo depositados en los centros de Archivo de sus organizaciones políticas o sindicales, como los Archivos de la Fundación José Barreiro donde se encuentran fondos de archivo orales de entrevistas realizadas a Purificación Tomás Vega (1918-1990, exilio 1937-1976), secretaria de la ejecutiva femenina de las Juventudes Socialistas. El fondo personal de Carmen García Bloise (1921-1990) está depositado en el Archivo de la Fundación Pablo Iglesias, y en el Archivo Histórico del Partido Comunista de España está depositado, desde 2013, el de Rosalía Sender Begué,

En el Archivo Histórico del Partido Comunista de España se deposita una serie denominada *Organizaciones de mujeres*, y los fondos documentales de los grupos de doble militancia: Movimiento Democrático de Mujeres, Movimiento de Liberación de la Mujer y Comisión de la Mujer del Comité Central del PCE. Hay depositados varios fondos de organizaciones anteriores al inicio de la Transición. Otro partido político, la Liga Comunista Revolucionaria mantiene actualmente un proyecto de digitalización a través de la Fundación Viento del Sur, denominado Historia de la LCR. Entre sus fondos se encuentra el feminismo en el proyecto político de la LCR. El Archivo Histórico de la Fundación 10 de Marzo de Santiago de Compostela, también llamado Archivo Histórico del Sindicato Nacional de Comisiones Obreras de Galicia, alberga fondos de Coordinadora Feminista Galega, Asociación Galega da Muller, Movimiento Democrático de Mujeres.

6. DISCUSIÓN

Como consecuencia de esta situación, es muy vulnerable la perdurabilidad de los fondos que albergan estos centros de colectivos feministas autónomos. Existen feministas que por su categoría y prestigio no tendrían problema en depositar, de forma individualizada, sus fondos personales en archivos históricos o universitarios. En cambio, por lo que respecta a la documentación del feminismo de doble militancia, existe mayor protección, al estar integrados en sistemas archivísticos ya organizados.

Los archivos universitarios consultados, por ejemplo, el Archivo General de la Universidad Complutense, el Archivo General de la Universitat d'Alacant y el Archivo General de la Universidad Carlos III, no excluyen en sus estatutos donaciones y cesiones que realicen otras personas físicas o jurídicas ajenas a los mismos. Se podría considerar esta posibilidad para salvaguardar el patrimonio documental del feminismo, sin embargo, esta solución ocasionaría una mayor dispersión y dificultad para realizar proyectos archivísticos generales.

A este respecto cabría hacer mención al proyecto del Grupo de Trabajo de Fondos Personales (GTFP) de la Conferencia de Archiveros de las Universidades Españolas, para tratar de unificar criterios en el cuadro de clasificación para los fondos personales de los distintos archivos universitarios (Conferencia de Archiveros de las Universidades Españolas, 2012). El trabajo continuado del GTFP está elaborando guías para el uso de cuadro de clasificación adoptado (Conferencia de Archiveros de las Universidades Españolas, 2013).

Los resultados de localización de fondos de esta temática tanto en centros de Archivo públicos como privados, revelan un porcentaje ínfimo recuperado. Sin embargo, las evidencias inducen a considerar un volumen enorme, y en el mejor de los casos, es de suponer no haya sido destruido o desaparecido. Tras el análisis, los datos resultan poco optimistas, se estima que un significativo caudal documental de la historia del movimiento feminista está en paradero desconocido. Puede hallarse parcelado entre los archivos administrativos de las entidades productoras que continúen en su actividad, los archivos personales de las feministas pertenecientes a dichos grupos ya desaparecidos, reinventados y otros paraderos. Además, de la posibilidad de que muchos de estos fondos se encuentren tanto en manos privadas como en grupos herederos, en todo caso, de tenencia todavía no identificada.

En cuanto a la desprotección legislativa mencionada en estos casos, invita a citar la idea expresada en foros profesionales (Heredia Herreros, 2012): "el retraso legislativo en archivística, ya que los plazos para formar parte del patrimonio protegido son de 40 y 100 años, y pudiera ocurrir que muchos no alcancen ese plazo por incuria, deterioro o destrucción".

La idea de proteger el patrimonio del feminismo español resulta tanto innovadora como necesaria; en los últimos años surgen algunas iniciativas en pro de la memoria histórica de las mujeres. Ejemplo representativo de ello es el proyecto "Álbum de Mulleres", de recuperación de memoria histórica de mujeres de Galicia. Cabe mencionar el proyecto de la Universitat de Barcelona "Diccionari Biogràfic de Dones" y el recientemente presentado "Museo de Hechos y Derechos de las Mujeres", para reunir y exponer documentos y materiales relacionados con la trayectoria histórica de los derechos de las mujeres, promovido éste por el feminismo académico de la Universitat d'Alacant (Asociación Museo de Hechos y Derechos de las Mujeres, 2014).

El resultado obtenido fundamenta la propuesta para tratar de salvaguardar el patrimonio documental, que pasa por plantear un sistema archi-

vístico y un centro de Archivo para albergar los fondos documentales del feminismo español. Un centro con capacidad de desarrollar una política archivística, necesaria para asegurar niveles óptimos de calidad, de rentabilidad y de eficacia (Cruz Mundet, 2006), precisos tanto para normalizar un sistema archivístico o reglamentario, como para la gestión documental y procedimental de un centro de estas características. Todo ello presupone abordar el estadio final de los fondos de archivo, ya sea este su conservación o su eliminación. Para asegurar criterios de organización originaria, sería adecuado que los agentes productores mantuviesen un modelo, formas y grado de organización (Núñez Fernández, 1999), con el centro destinatario final de los documentos.

El futuro Centro de Archivo del Feminismo, del cual se propone su creación, ha de disponer de un sistema de archivo que garantice a largo plazo el mantenimiento, la conservación de la autenticidad, fiabilidad y accesibilidad de los documentos depositados, utilizando el estándar ISO 154891 y 2, para gestión de archivos (ISO, 2001). Una vez consolidada una primera fase del proyecto de localización y depósito de fondos, se ha de concatenar una parte técnica de digitalización, y un sistema de gestión documental digital, diseñado con los requerimientos funcionales esenciales siguiendo las especificaciones de *MoReq2* para la administración de los archivos electrónicos.

Se recomienda impulsar un sistema de Archivo, como una ordenación y puesta a disposición de los materiales para el personal investigador, habilitando un centro no solo para salvaguardar, sino también para dar servicio de Archivo en condiciones adecuadas. Destacan los proyectos de sistemas de Archivo privados realizados en España a partir de la Transición democrática y que perduran en el tiempo, como el Centro de Archivo de la Fundación Francisco Largo Caballero, el Archivo de la Fundación Pablo Iglesias, y el Archivo Histórico del Partido Comunista de España. Iniciados los tres a finales de la década de los años 70, cumplieron con el objetivo de recuperar y difundir el patrimonio documental propio hasta la actualidad (Archivos para la historia del movimiento obrero español, 1985).

Para la protección del patrimonio y la preservación de los documentos de archivo para la investigación histórica, no es suficiente conocer las fuentes escritas gracias al esfuerzo personal de unas cuantas militantes feministas, sino que se hace necesaria la creación de un organismo que articule todo este caudal documental que, mayormente está inédito y disperso. También es muy relevante velar por el acceso a la información de la ciudadana,

y considerar en el desarrollo del proyecto la importancia y la responsabilidad del Estado en el papel facilitador de la preservación, organización y acceso a los archivos privados de interés social (Porto Atona López, 2011).

En las universidades españolas, mencionaba Torres Ramírez (2000): "es muy importante la tarea investigadora realizada por quienes cultivan los Estudios de las Mujeres". Se considera que hay dos instituciones de apoyo constituidas que representan el feminismo académico estatal. De una parte, la Asociación Española de Investigación de Historia de las Mujeres (AEIHM), que coordina la labor desarrollada en los distintos Seminarios de Estudios de las Mujeres existentes en las Universidades y Centros de Investigación del Estado Español, encargada de fomentar la investigación y divulgar los estudios feministas y de Historia de las Mujeres a escala nacional e internacional, así como promover la investigación en el campo de la Historia de las Mujeres y políticas de Género. Y por otra, la asociación de ámbito académico AUDEM que coordina los Institutos, Seminarios y Centros de Estudios de las Mujeres o de Género existentes, o que puedan existir en el futuro, en las distintas Universidades del Estado español. Fomentando los Estudios de las Mujeres y de Género a todos los niveles de la docencia e investigación.

El motor primigenio para realizar un proyecto de tal envergadura es el asociacionismo del feminismo, pero han de ser las políticas públicas las que lo materialicen. Ante la ausencia de estas políticas actualmente se está gestando la *Asociación para la Protección del Patrimonio Documental Feminista*, un proyecto asociativo inclusivo del feminismo académico representado por los institutos y seminarios de investigación feminista de las universidades, los colectivos feministas y los centros de documentación y bibliotecas de mujeres de España. La asociación servirá de puente entre las personas interesadas en garantizar y preservar los fondos documentales de las asociaciones y personas vinculadas al feminismo.

7. CONCLUSIONES

Tras el análisis de la situación son numerosas las razones para proyectar el Centro de Archivo del Feminismo. Otra de las razones que contribuyen a que las mujeres se encuentren menos representadas en los libros de historia es la precariedad de las fuentes primarias para el estudio de la historia y la memoria, "los archivos". Las mujeres en general y las activistas feministas en particular deberían conocer que sus documentos privados y sus relatos de vida son valiosos para la historia y la memoria colectiva.

Por lo que respecta a la localización de los fondos documentales de archivo del feminismo de la Transición democrática, según los resultados observados, ante la carencia de un centro de Archivo para depositar los fondos, y considerando el volumen potencial de los grupos y asociaciones de mujeres del Movimiento Feminista, se concluye que existe una gran pérdida documental, que en su mayoría se encuentra en paradero aún desconocido.

Se manifiesta la necesidad de creación de un sistema archivístico de los fondos documentales y de un Centro de Archivo para salvaguardar el patrimonio documental del feminismo. Tarea pendiente de organización que ha de materializarse de forma similar a como se ha realizado en proyectos semejantes en el sindicalismo obrero y otros movimientos sociales a finales de la década de los años 70.

En cuanto a poner en valor el patrimonio documental del feminismo, se constata la necesidad de sensibilización, fomentando redes que incluyan todas las entidades involucradas: los colectivos feministas, las instituciones universitarias que investigan y los centros de documentación y bibliotecas de mujeres, posibilitando acciones comunes.

La creación del Centro de Archivo del Feminismo, que asegure la preservación de los documentos

para la historia y la investigación, que resolvería el problema que ocasiona la recopilación, tratamiento, depósito documental y otros muchos aspectos de archivos privados feministas.

Para facilitar el estudio de estas fuentes documentales, también es necesaria la creación de una guía de fuentes para los estudios de género y la historia del movimiento feminista que incluya la descripción de todos los fondos de los archivos identificados.

Necesariamente la historia ha de incluir la vida y acción de las mujeres y su legado. Nos decía Octavio Paz que la poesía es la memoria de la vida y los archivos son su lengua. No podemos seguir arrinconando la historia de las mujeres y la memoria de sus vidas. Las mujeres no estamos excluidas en la historia, sino incluidas como ausencia.

8. AGRADECIMIENTOS

Las autoras agradecen la colaboración en la recopilación de los datos a Marisa Mediavilla de la Biblioteca de Mujeres de Madrid y a los centros de documentación y bibliotecas de mujeres. Asimismo extendemos el agradecimiento a los revisores anónimos que con sus comentarios y sugerencias han contribuido a la mejora sustancial de este trabajo.

9. REFERENCIAS

- Acta de la Reunión celebrada el día 9 de octubre de 1989 entre representantes del Instituto de la Mujer y la Comisión consultiva para el seguimiento del Plan para la Igualdad de Oportunidades de las Mujeres. (1988-1990). Archivo personal de Ana M^a Pérez del Campo Noriega. C. 12. Cp. 52.
- Alberch i Fugueras, R.; Boix Llonch, L.; Navarro Sastre, N.; Vela Palomares, S. (2001). *Archivos y cultura: manual de dinamización*. Gijón: Trea, p. 173.
- Anderson, B. (1991). *Historia de las mujeres: una historia propia*. Barcelona: Crítica.
- Archivos para la historia del movimiento obrero español: XV Conferencia de la IALHI, 25-28 de septiembre de 1984*, Madrid: Fundación Pablo Iglesias, 1985.
- Asociación "Mujeres en la Transición Española" (1999). *Españolas en la Transición: de excluidas a protagonistas (1973-1982)*. Madrid: Biblioteca Nueva.
- Asociación Museo de Hechos y Derechos de las Mujeres (2014). *Jornada Feminista. "Ayer y hoy del feminismo: de las olas a las mareas"* [Videograbación]. [Disponible en: <http://www.youtube.com/watch?v=J6AY4vipueU>]. [Consulta: 12/11/2014].
- Augustín Puerta, M. (2003). *Feminismo, identidad personal y lucha colectiva: (análisis del movimiento feminista español en los años 1975 a 1985)*. Granada: Universidad de Granada.
- Caballé Masforroll, A. (2006). El bolso de Ana Karenina. La necesidad de inventariar los textos autobiográficos. En: *Seminario de archivos personales (Madrid, 26 a 28 de mayo de 2004)*. Madrid: Biblioteca Nacional, p. 195-210.
- Capel Martínez. R. M. (2009). El archivo de la Residencia de Señoritas. *Participación Educativa*. Madrid: Consejo Escolar de Estado, 11, (julio), p. 156-161. [Disponible en: <http://www.mecd.gob.es/revista-cee/pdf/n11-capel-martinez.pdf>]. [Consulta: 04/10/2014].
- Codina_Canet, M. A. (2013). *Biblioteca de Educación: Historia de una institución*. [Madrid]: Ministerio de Educación, Cultura y Deporte, Subdirección General de Documentación y Publicaciones, p. 12.
- Los Colectivos Feministas se definen (1976). *Vindicación Feminista*, 3, p.15-26.
- Conferencia de Archiveros de las Universidades Españolas. GTFP (2012). Propuesta de cuadro de clasificación común de fondos personales (v.2.0). *XVIII Jornadas*

- de la Conferencia de Archiveros de las Universidades Españolas. Universidad de Cádiz, 23, 24 y 25 de mayo de 2012. [Disponible en: http://cau.crue.org/Documents/GT/GT-AP/XVIII_Jornadas_CAU_UCA_Cadiz_GT_Fondos_Personales_25_05_2012.pdf]. [Consulta: 29/09/2014].
- Conferencia de Archiveros de las Universidades Españolas. GTFP (2013). *Informe de actividades*. 2012 Informe elaborado a solicitud del Comité Ejecutivo CAU 27 - febrero de 2013. Grupo de Trabajo de Fondos Personales de las actividades desarrolladas en 2012. [Disponible en: http://cau.crue.org/Documents/GT/GT-AP/Informe_FONDOSPERSONALES_2014.PDF]. [Consulta: 29/09/2014].
- Cortés Alonso, V. (1986). El Concepto de Archivo en Pardo Bazán y algunas reflexiones. *Boletín Auriense / Museo Arqueológico Provincial*, 14-15, p. 173-186.
- Cruz Mundet, J. R. (2006). *La gestión de documentos en las organizaciones*. Madrid: Ediciones Pirámide, 2006.
- Cuesta Bustillo, J. (2003). Mujeres y Democracia en España (1975-2000). En: *Historia de las Mujeres en España siglo XX*. Madrid: Instituto de la Mujer.
- Elvira i Silleras, M. (2013). Posibilidades de eliminar documentos en archivos personales. En: *VI Jornadas Archivando: la valoración documental*. León, 7 y 8 de noviembre 2013. p. 219-230. [Disponible en: https://archivosierrapambley.files.wordpress.com/2014/02/maria_elvira.pdf]. [Consulta: 6/12/2014].
- Folguera Crespo, P. (1997). Democracia y cambio social. En: Garrido González, Elisa. *Historia de las mujeres en España*. Madrid: Síntesis.
- Fraile, M. A. (2000). La asociación 'María Moliner' historia de una asociación de bibliotecas de mujeres. En: *Métodos de Información*, 7 (35-36), p. 81-84. [Disponible en: <http://www.metodosdeinformacion.es/mei/index.php/mei/article/viewFile/485/505>]. [Consulta: 18/08/2014].
- Guía de asociaciones de Mujeres* (1996). Madrid: Consejo de la Mujer.
- Guía de asociaciones de mujeres y centros de interés, 1989* (1989). Madrid: Ministerio de Asuntos Sociales.
- Guía de las asociaciones de mujeres* (1987). Madrid: Instituto de la Mujer.
- Heredia Herreros, A. (2012). Legislar sobre archivos. Reflexiones y experiencias. *III Foro de legislación Archivística. Conmemorativo de los 150 años de la Batalla del 5 de mayo de Puebla y a la Revolución Mexicana*. Universidad Autónoma de Puebla (México), 30 de noviembre de 2012 [Videograbación]. [Disponible en: <https://www.youtube.com/watch?v=Ipn8iPaSWho>] [Consulta: 06/12/2014].
- Herrero Montero A. M.; Díaz Rodríguez, A. (2011). La Clasificación archivística. Cruz Mundet, J. R. (ed.). *Administración de documentos y archivos. Textos fundamentales* [en línea]. Madrid, Coordinadora de Asociaciones de Archiveros, Ministerio de Cultura. [Disponible en: <http://www.archiveros.net/LIBRO.ARCHIVOS.IBEROAMERICANOS.pdf>]. [Consulta: 6/12/2014].
- International Standards Organization (2001). Information and documentation. Records management. ISO 15489 1 Y 2. 2001 Part 1: General y Part 2: Guía de aplicación. ISO. Genève, 2001. [6. Política y responsabilidades.6.1. Aspectos generales].
- Martínez Ten, C.; Gutiérrez López, P.; González Ruiz, P. (2009). *El movimiento feminista en España en los años 70*. Madrid: Cátedra: Instituto de la Mujer; Valencia: Universitat de València.
- Mastropiero, M. C. (2006). Archivos privados: análisis y gestión. Buenos Aires: Alfagrama.
- Muñoz-Muñoz, A. M.; Argente Jiménez, M. (2010). Red de centros de documentación y bibliotecas de mujeres: cooperación entre las bibliotecas feministas españolas. *El profesional de la información*, 19 (5), p. 505.
- Nash, M. (2007). *Dones en transició: de la resistència política a la legitimitat feminista: les dones en la Barcelona de la Transició*. Barcelona: Ajuntament, Regiduria de Dona.
- Nash, M. (1994). Experiencia y aprendizaje: la formación histórica de los feminismos en España. *Historia Social*, 20, p. 151-172.
- Núñez Fernández, E. (1999). *Organización y gestión de archivos*. Gijón: Trea.
- Pazos, M. (2014). *Homenaje a María Telo (1915-2014: In memoriam)*. Madrid: Plataforma por Permisos Iguales e Intransferibles de Nacimiento y Adopción.
- Pérez Pais, C.; Mariño Costales, M. (2013). *Recuperación da documentación e memoria do movemento feminista organizado en Galicia*. Santiago de Compostela: Consello da Cultura Gallega. [Disponible en: http://consellodacultura.org/mediateca/extras/329a23_Movemento%20Feminista.pdf]. [Consulta: 12/11/2014].
- Porto Ancona López, A. (2011). Archivos y ciudadanía: el acceso a la información pública. *Revista General de Información y Documentación*, 21, p. 249-264. [Disponible en: <http://revistas.ucm.es/index.php/RGID/article/view/249-264>]. [Consulta: 25/01/2015].
- Rubio, O. M. (2012). *100 años en femenino: una historia de las mujeres en España: ... Centro Conde Duque, entre el 9 de marzo y el 20 de junio de 2012*. [Madrid]: Acción Cultural Española.
- Suárez Suárez, C. (2005). Feministas en la transición asturiana. *La aljaba*, 9 (ene./dic.). [Disponible en: <http://www.scielo.org.ar/scielo.php?pid=S1669->

- 57042005000100002&script=sci_arttext]. [Consulta: 6/12/2014].
- Suárez Suárez, C. (2012) *El feminismo asturiano en la oposición al Franquismo y en la Transición Democrática. Vivencias, conciencia y acción política*. Tesis doctoral. Universidad de Oviedo.
- Subvenciones a asociaciones de mujeres y ayuntamientos 1990 (1990). *Revista 8 de Marzo*, 1, p. 27.
- Telo Núñez, M. (2009). *Mi lucha por la igualdad jurídica de la mujer*. Cizur Menor, Navarra: Aranzadi.
- Telo Núñez, M. (1998). Las investigadoras de la mujer. *ABC*, de 18 de enero de 1998. [Disponible en: <http://hemeroteca.abc.es/nav/Navigate.exe/hemeroteca/madrid/abc/1998/01/18/064.html>]. [Consulta: 03/11/2014].
- Torres Ramírez, I. de (2000). El Feminismo académico en España hoy. *Métodos de Información*, 7 (35-36), p. 66. [Disponible en: <http://www.metodosdeinformacion.es/mei/index.php/mei/article/viewFile/481/501>]. [Consulta: 24/10/2014].
- Vázquez Murillo, M. (2006). *Cómo seleccionar documentos de archivo*. Buenos Aires: Alfagrama.



ESTUDIOS / RESEARCH STUDIES

Factores determinantes en la producción científica de los grupos de investigación en Colombia

Gladys Rueda-Barrios*, Manuel Rodenes-Adam**

* Grupo de Investigación en Administración, Facultad de Administración de Empresas, Universidad Pontificia Bolivariana, Edificio E- 202, Bucaramanga, Colombia.

** Grupo de Investigación ITIO, Departamento de organización de Empresas, Universidad Politécnica de Valencia. Correo-e: gladys.rueda@upb.edu.co ; mrodenes@omp.upv.es

Recibido: 26-02-2015; 2ª versión: 21-04-2015; Aceptado: 05-05-2015.

Cómo citar este artículo/Citation: Rueda-Barrios, G. y Rodenes-Adam, M. (2016). Factores determinantes en la producción científica de los grupos de investigación en Colombia. *Revista Española de Documentación Científica*, 39(1): e118. doi: <http://dx.doi.org/10.3989/redc.2016.1.1198>

Resumen: Existe un gran debate sobre la forma como deben estar organizadas las instituciones creadoras de conocimiento, es el caso de las universidades, grupos y centros de investigación; así como, los procesos claves de gestión del conocimiento que deben aplicar los grupos de investigación; y los recursos financieros, humanos y tecnológicos necesarios para lograr un alto rendimiento en la investigación. El presente trabajo pretende, a partir de la revisión teórica y el estudio empírico, determinar la relación entre los resultados de producción científica obtenidos por los grupos de investigación adscritos a universidades en Colombia, respecto a las variables cultura organizacional, gestión del conocimiento y capital tecnológico. Para el análisis, se obtuvo una muestra de 223 grupos de investigación a nivel nacional, y que se encuentran registrados y categorizados por el Departamento Administrativo Nacional de Ciencia, Tecnología e Innovación en Colombia (Colciencias, 2010). Los datos analizados se obtuvieron a partir de la aplicación de una encuesta que permitía medir la percepción de los líderes investigadores sobre las variables independientes cultura de la universidad, procesos de gestión de conocimiento y capital tecnológico; y la variable dependiente, producción científica, se analizó a partir de la calificación que asigna Colciencias a cada grupo por los resultados de investigación obtenidos durante un periodo de tiempo. Las técnicas estadísticas utilizadas en este estudio son: el análisis de regresión y el análisis de ecuaciones estructurales, con los cuales se determinó que la cultura motivadora, el proceso de externalización en la gestión del conocimiento y el capital tecnológico influyen de forma positiva en los resultados obtenidos en producción científica.

Palabras clave: Cultura organizacional; gestión del conocimiento; capital tecnológico; producción científica; grupos de investigación de Colombia.

Determining factors in Colombian research groups' scholarly output

Abstract: There is much debate about how knowledge-creating institutions —such as universities, research groups and centers— should be organized, as well as about the key processes of knowledge management to be applied by research groups and the financial, human and technological resources needed to achieve high performance in research. The current work, based on a theoretical review and empirical study, aims to determine the relationship of scholarly production results obtained by research groups affiliated with universities in Colombia in terms of the variables of organizational culture, knowledge management, and technological capital. For this analysis, a sample was obtained of 223 research groups that are recognized and categorized by the National Administrative Department of Science, Technology and Innovation (Colciencias, 2010). The data analyzed were obtained from a survey that measured the perception of leading researchers about such independent variables as university culture, knowledge management processes, and technological capital. The dependent variable of scientific production was analyzed based on the rating assigned to each group by Colciencias for research results obtained over a period of time. The statistical techniques used in this study are regression analysis and structural equation analysis by which it was determined that the following variables have a positive influence on the results of scholarly output: the motivating culture, process of outsourcing knowledge management, and technological capital.

Keywords: Organizational culture; knowledge management; technological capital; scientific production; Colombian research groups.

Copyright: © 2016 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0.

1. INTRODUCCIÓN

La investigación es un pilar fundamental dentro de la misión universitaria, que contribuye al desarrollo económico, político y social de los países; representa también un valor agregado para la imagen y prestigio de las universidades de acuerdo a los resultados de producción científica alcanzados. Por este motivo, la universidad y los grupos de investigación son objeto de importantes análisis y reflexiones desde diferentes disciplinas, tales como la psicología, sociología, economía, educación, administración, ingenierías y científicos en general.

Los grupos, centros e institutos de investigación adscritos a las universidades son los encargados de producir conocimiento, estudiar distintos fenómenos, ayudar a resolver problemas de las empresas e impactar de forma positiva en todos los campos de la sociedad y del conocimiento; por lo tanto, al resultado de estas investigaciones al hacerse público se le denomina producción científica; convirtiéndose en un tema de especial interés desde finales del siglo XIX por las universidades alemanas, quienes empezaron a considerarla como una nueva actividad que permitía ser referente para las demás universidades europeas y norteamericanas (Clark, 1997).

En Colombia, el ejercicio de la investigación se desarrolla principalmente a través de grupos de investigación promovidos por las universidades y el Departamento Administrativo Nacional de Ciencia, Tecnología e Innovación – Colciencias. Se denomina grupo de investigación "*al conjunto de personas que interactúan para investigar y generar productos de conocimiento en uno o varios temas, de acuerdo con un plan de trabajo de corto, mediano o largo plazo (tendiente a la solución de un problema). Un grupo es reconocido como tal, siempre que demuestre continuamente resultados verificables, derivados de proyectos y de otras actividades procedentes de su plan de trabajo*" (Colciencias, 2010); actualmente, Colciencias reconoce 4.072 grupos de investigación, de los cuales 3.653 pertenecen a universidades y 71 son centros de investigación. Los resultados de investigación de los grupos están condicionados por la influencia de muchas variables, entre ellas la cultura de las instituciones universitarias, los procesos internos de gestión de conocimiento en los grupos y las instituciones, el apoyo económico para la investigación, el interés por parte de sus directivos para lograr avances y los recursos tecnológicos, entre otros. La investigación es un proceso que requiere de una serie de insumos para lograr los resultados esperados, los cuales deben producir unos output o salidas denominados productos, tales como artículos científicos publicados en revistas de alto impacto,

libros, patentes, ponencias, entre otros (Bermeo, 2007; Rivera y otros, 2009).

Desde la perspectiva de la producción científica, los estudios relacionados con las entradas o insumos se han centrado en diferentes unidades de análisis, algunos realizados a las instituciones académicas, otros centrados en los grupos de investigación, áreas de investigación o investigadores individuales, por lo que se recomienda que las evaluaciones de rendimiento y eficiencia en la investigación deben realizarse a nivel de grupos o "micro-unidades" y a nivel global o "macro-unidades", porque de esta manera se logra un análisis detallado que puede proporcionar conclusiones interesantes (Cherchye y Vanden Abeele, 2005).

Entre los factores explicativos e indicadores asociados a la producción científica se encuentran los factores institucionales, culturales, financieros, profesionales, atributos personales y demográficos. Sin embargo, hasta el momento no hay un criterio único que permita establecer y estandarizar los determinantes de la producción científica para lograr un alto rendimiento en la investigación (Avital y Collopy, 2001).

Basados en la literatura científica (Avital y Collopy, 2001; Bozeman y Corley, 2004; Rivera y otros, 2009; Heinze y otros, 2009) se propone para esta investigación un modelo de investigación aplicado a los grupos de investigación de las universidades de Colombia, que incluye una variable dependiente denominada producción científica y tres variables independientes consideradas factores determinantes: la cultura de las instituciones universitarias, los procesos de gestión de conocimiento y el capital tecnológico.

Un primer punto clave en este trabajo, es identificar la cultura organizacional y las dimensiones culturales que intervienen en el desarrollo de las actividades de investigación de las universidades. Algunos estudios han analizado a las universidades, su forma de organización y las estructuras necesarias que deben mantener en estos tiempos de constantes cambios para dar lugar a una mejor labor en la docencia, la investigación y la extensión (Clark, 1997; Harvey y otros, 2002; Van der Meulen, 2002; Czerniewicz y Brown, 2009; Rodríguez y Páez, 2009). En este sentido, se entiende que las universidades son organizaciones con recursos y procesos que buscan ofrecer servicios en el sector de la educación y por lo tanto poseen una cultura organizacional propia que las identifica y que interviene en el desarrollo de las actividades diarias, en el quehacer científico y en la generación del conocimiento; ese carácter particular hace que se puedan citar distintos tipos de universidades, con distintos valores, prácticas y estructuras.

Una segunda variable, reconocida por muchas organizaciones hoy en día, es la gestión del conocimiento. Las universidades son organizaciones creadoras de conocimiento y por lo tanto se consideran actores principales en la creación y difusión del conocimiento a la sociedad; a ellas están adscritos los grupos, centros y/o institutos que trabajan en investigación en Colombia. En consecuencia, es importante el análisis de los procesos de la gestión del conocimiento en estas instituciones, particularmente en los grupos de investigación, logrando un mayor protagonismo de la investigación universitaria en los sectores empresariales, sociales y estatales (Nonaka y Takeuchi, 1995; Bueno y otros, 2003; Jaime y otros, 2005; Gaviria y otros, 2007).

La última variable independiente del modelo tiene en cuenta el capital tecnológico de los grupos de investigación que permite dar soporte a las labores investigativas, tales como recursos bibliográficos, laboratorios, equipos, licencias, personal I+D, entre otros; además de considerar los recursos basados en las tecnologías que están lideradas por internet y otras tecnologías emergentes que han impactado en el desarrollo de la producción científica. En el caso de las universidades, y propiamente en los grupos de investigación, además de las tecnologías que puedan incorporar en sus procesos básicos de investigación, también el campo virtual se ha convertido en una herramienta clave ya que permite el desarrollo interno de la investigación propia del grupo, pero además permite el intercambio de conocimiento, la colaboración con grupos externos a la institución, fortaleciendo su propia investigación y aumentando el logro de resultados.

Las tres variables seleccionadas han sido analizadas por otras investigaciones de forma independiente (Bueno y otros, 2003; Jaime y otros, 2005; Acevedo y otros, 2005; Harvey y otros, 2002; Jaime y Blanco, 2007; Gaviria y otros, 2007; Rodríguez y Páez, 2009), ya que su relación con la producción científica puede tener una incidencia fuerte; y por este motivo el presente trabajo quiere comprobar la existencia de la relación de estas variables independientes con la producción científica, y además la relación existente entre sí de las variables independientes.

2. OBJETIVOS DE ESTA INVESTIGACIÓN

El objetivo general es analizar la cultura de la organización, los procesos claves de la gestión del conocimiento, el capital tecnológico, y su relación con los resultados de producción científica obtenidos por los grupos de investigación adscritos a universidades en Colombia.

Los objetivos específicos que se desprenden del objetivo general son:

- Determinar la relación entre la cultura de la universidad y los resultados de producción científica.
- Identificar los procesos de gestión del conocimiento de los grupos de investigación basados en el Modelo SECI¹ y establecer su relación con los resultados de producción científica.
- Identificar el capital tecnológico de los grupos de investigación y establecer su relación con los resultados de producción científica.
- Validar el modelo teórico mediante el análisis de resultados obtenidos del sector objeto de estudio considerando los grupos de investigación de universidades en Colombia.

3. MODELO E HIPÓTESIS DE ESTA INVESTIGACIÓN

Con base en el planteamiento del problema y la revisión teórica, se incluyen a continuación las hipótesis:

Hipótesis 1: *La cultura organizacional de una universidad basada en valores y prácticas orientados a la participación, el trabajo en equipo, la interdisciplinariedad, la innovación y el interés por la investigación, se relaciona positivamente con los resultados de Producción Científica.*

En el siglo XXI, las universidades deben tener una transición de valores y prácticas tradicionales hacia la evolución de nuevos esquemas culturales que propicien nuevas políticas universitarias para los procesos de investigación, las prácticas de los profesores y la manifestación de estos mediante los resultados de alta calidad en la docencia y la investigación. En este sentido, esta investigación pretende identificar factores críticos de la cultura de las instituciones universitarias y su relación con los resultados de producción científica obtenidos por los grupos de investigación (Clark, 1997; Avital y Collopy, 2001; Harvey y otros, 2002; Van Der Meulen, 2002; Gaviria y otros, 2007; Rodríguez y Páez, 2009; Tomas y Rodríguez, 2009).

Hipótesis 2: *Los procesos claves de gestión del conocimiento que involucra cada una de las actividades del modelo SECI, se relacionan positivamente con los resultados de Producción Científica.*

Se han realizado estudios que permiten medir los procesos de gestión del conocimiento en los centros de investigación. Citando algunos de los resultados encontrados, se evidencia que los procesos claves que permiten la gestión del conocimiento como la socialización, exteriorización, combinación

e internalización propician el aumento de nuevos conocimientos (Arbonés, 1999; Bueno y otros, 2003; Jaime y otros, 2005; Jaime y Blanco, 2007; Gaviria y otros, 2007). En este sentido existen prácticas y rutinas relacionadas con la compartición y divulgación de conocimientos que llevan a cabo los grupos de investigación y que influyen en una mayor producción científica.

Hipótesis 3: *El Capital Tecnológico que involucra esfuerzo I+D, personal y tiempo I+D, dotación tecnológica, y uso de las TIC para la investigación, se relacionan positivamente con los resultados obtenidos en Producción Científica.*

La capacidad que tienen los grupos de investigación en cuanto a dotación tecnológica, las tecnologías de información, inversión en I+D e infraestructura permiten pronosticar mejores resultados de producción científica (Acevedo y otros, 2005). El hecho de que un grupo de investigación tenga las herramientas y recursos necesarios propicia mayor investigación y una mayor motivación por parte de los investigadores, por lo tanto se reconoce la importancia del capital tecnológico en los procesos de investigación y en los resultados obtenidos (Bueno y otros, 2003).

Hipótesis 4: *La cultura organizacional de una universidad basada en valores y prácticas orientados a la innovación, la participación, el trabajo en equipo, la interdisciplinariedad e interés por la investigación, se relaciona positivamente con los procesos de Gestión del Conocimiento.*

La cultura es un factor que interviene en las actividades diarias de la gestión del conocimiento y por lo tanto en este caso de las actividades de los docentes e investigadores, por ello en la literatura se propone un nuevo esquema de Organización denominado Hipertexto (Nonaka y Takeuchi, 1995), en el que los procesos de gestión del conocimiento se llevan a cabo basados en la cultura de la organización y en las características culturales que involucran políticas, valores y prácticas de una organización creadora de conocimiento. De acuerdo al tipo de cultura en una organización, se puede facilitar o no el proceso de gestión del conocimiento; aspectos como los incentivos, espacios de investigación, trabajo interdisciplinario, tiempo, autonomía, participación y carácter institucional pueden redundar en la creación y conversión del conocimiento.

Hipótesis 5: *La cultura organizacional de una universidad basada en valores y prácticas orientados a la innovación, la participación, el trabajo en equipo, la asignación de recursos, la adopción y uso de tecnologías, los esfuerzos y recursos tecnológicos para la investigación se relacionan positivamente con el Capital Tecnológico.*

La cultura organizacional emprendedora y profesional contempla valores orientados a la importancia de la formación, la divulgación de eventos científicos, asignación de recursos, promoción de alianzas y convenios, y el uso de nuevas tecnologías, influyendo todas ellas en el capital tecnológico, tales como, mayores recursos de I+D, personal I+D, dotación tecnológica y acceso a nuevas tecnologías de la información y la comunicación (Bueno y otros, 2003; Lopez-Nicolas y Meroño, 2009).

Hipótesis 6: *El Capital Tecnológico basado en el esfuerzo I+D, personal y tiempo I+D, dotación tecnológica, y uso de las TIC para la investigación se relaciona positivamente con los procesos de Gestión del Conocimiento.*

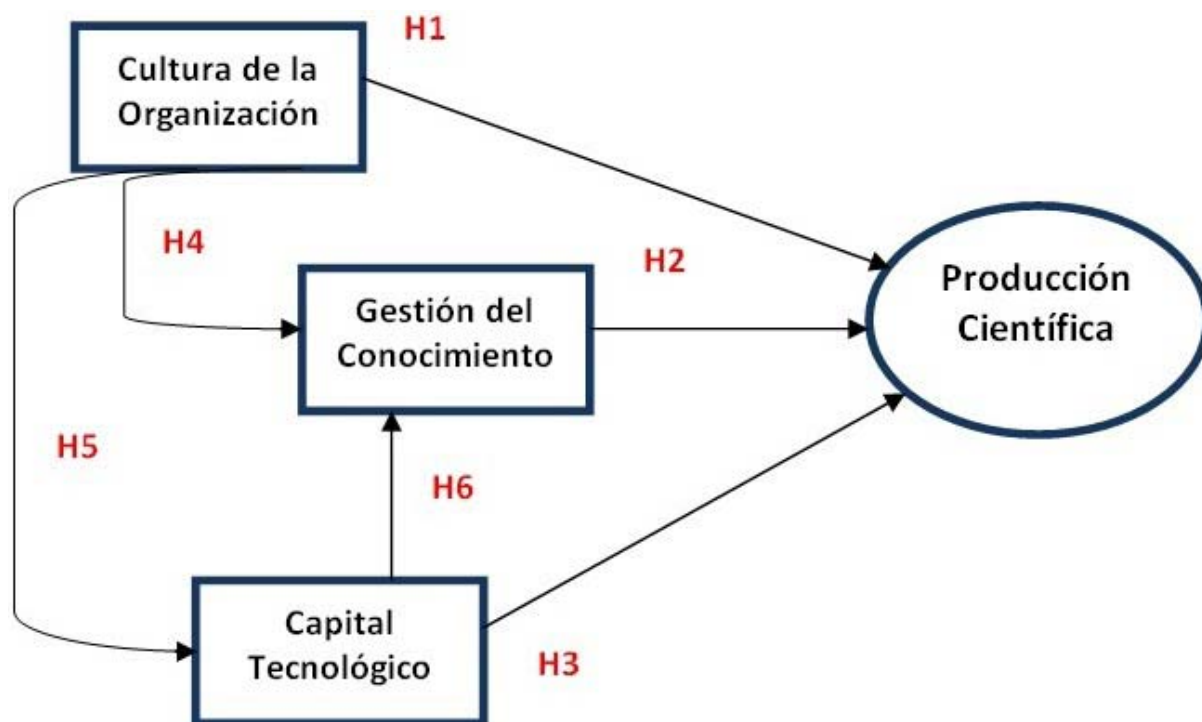
En el análisis teórico se encuentra una relación entre la gestión del conocimiento y el capital tecnológico; es importante contar con los recursos tecnológicos necesarios para el desarrollo de la investigación y que son útiles en cada uno de los procesos de gestión del conocimiento (Bueno y otros, 2003). Es decir, se considera necesario para que los procesos de gestión del conocimiento sean más rápidos y adecuados contar con los recursos tecnológicos relacionados con las tecnologías de la información (Chang y otros, 2005; Meroño, 2005; Lopez-Nicolas y Meroño 2009).

Según el planteamiento y la fundamentación de las hipótesis, en la figura 1 se visualiza el modelo de investigación y las relaciones planteadas entre las variables.

3.1 Metodología

Este trabajo tiene un alcance descriptivo, correlacional y explicativo; utilizando el método deductivo de manera que permite conocer la realidad de forma empírica; utilizando como herramienta para la recolección de datos un instrumento cuantificable (encuesta) aplicado a los grupos de investigación adscritos a las universidades en Colombia en un único periodo de 4 meses en el año 2011.

El número total de grupos de investigación clasificados por Colciencias es de 4.072 (incluye grupos de empresas, universidades y centros oficiales de investigación), distribuidos en la categoría A1 (187), A (256), B (652), C (933), D (2.044). De éstos, se consideraron como población objeto para el estudio 3.653 grupos de investigación adscritos exclusivamente a las universidades, de acuerdo con la información obtenida en la plataforma ScienTI-Col, en la última convocatoria de medición de grupos (Colciencias, 2010).

Figura 1. Modelo de investigación y relaciones entre variables

Para calcular el tamaño de la muestra se aplicó la fórmula de muestreo aleatorio simple (Miquel, 1997), con un nivel de confianza del 95% y $z=1,96$, una variabilidad positiva de $p=0,5$ y negativa $1-q=0,5$ y considerando un error del 5%, el tamaño de la muestra es de 347 casos (grupos de investigación), de los que se obtuvo respuesta de 223 casos, por lo que el error de estimación que se puede estar cometiendo es de 6,35%.

El estudio consideró como variable dependiente el indicador de producción científica de cada grupo, según el escalafón de grupos de investigación creado y clasificado por Colciencias (2008), que se denomina índice ScientiCol y que se encuentra visible en la Plataforma virtual ScienTI-COL. El proceso consiste en hacer una ponderación de los productos resultados de investigación del grupo en un período de observación (últimos 5 años), teniendo en cuenta el número de integrantes del grupo y la edad del grupo; es decir, un grupo tiene una puntuación base de acuerdo con una condición mínima de entrada al modelo y aumentará su condición de acuerdo a los requerimientos de existencia, calidad y visibilidad de la producción, como se puede ver en la Tabla I.

TABLA I. Categorías de los grupos de investigación

Categoría	ScientiCol
A1	<ul style="list-style-type: none"> ScientiCol mayor o igual a: 9,0 Edad del grupo mayor o igual a 5 años
A	<ul style="list-style-type: none"> ScientiCol mayor o igual a: 7,0 Edad del grupo mayor o igual a 5 años
B	<ul style="list-style-type: none"> ScientiCol mayor o igual a: 4,0 Edad del grupo mayor o igual a 3 años
C	<ul style="list-style-type: none"> ScientiCol mayor o igual a: 2,0 Edad del grupo mayor o igual a 2 años
D	<ul style="list-style-type: none"> ScientiCol mayor o igual a: 0,0 Edad del grupo mayor o igual a 1 año

Fuente: Colciencias, 2010.

Las condiciones mínimas para la categorización de un grupo exigen:

- Estar registrado en el sistema GrupLac de la Plataforma ScienTI - Colombia en Colciencias.
- Tener uno o más años de existencia.
- Estar avalado al menos por una (1) Institución registrada en el sistema InstituLAC de la Plataforma ScienTI - Colombia, a la cual el grupo haya registrado que pertenece.

- Tener al menos un proyecto de investigación en marcha.
- Que por lo menos una (1) de las personas vinculadas al grupo como investigador, posea una formación de pregrado, maestría, o doctorado concluida.
- Tener registrado al menos un (1) producto de nuevo conocimiento desarrollado por el grupo de investigación obtenido dentro de la ventana de observación.
- Reportar al menos dos productos resultantes de actividades de investigación relacionadas con la formación y la apropiación social del conocimiento, divulgación, extensión, o una combinación de éstas.
- Todos los productos resultados del quehacer científico de los miembros del grupo de investigación en un período de observación (últimos 5 años) deben estar registrados en la plataforma virtual ScienTI-COL; así mismo, se deben adjuntar los archivos que contienen los documentos requeridos en cada uno de los productos, necesarios verificar su existencia.

El índice ScientiCol se calcula asignando un peso según las características de cada producto (entre 0-1) y agrupando los productos según su tipología, de esta manera se crea un indicador que es dividido en un umbral definido para obtener los índices, como se muestra en la siguiente fórmula (Colciencias, 2008):

$$\text{ScientiCol} = 5 * (\text{NC}) + 3,5 * (\text{NCA}) + 1 * (\text{F}) + 0,5 * (\text{D})$$

y que se expresa así:

Índice de producción de Nuevo Conocimiento (NC): Este indicador está conformado por un subconjunto de productos, tales como artículos, libros, capítulos de libros, patentes, productos registrados, empresas spin off, normas. Los pesos asignados para el indicador de existencia oscilan entre 0 y 1; en el indicador de calidad el peso oscila entre 3,0 y 0,1 dependiendo del tipo de producto y el subtipo que indica la relevancia o impacto de la publicación, la editorial o revista indexada o reconocida.

$\text{INC} = \text{Art. A} + \text{Art. B} + \text{Art. C} + \text{Art. O} + \text{Libro A1} + \text{Libro A2} + \text{Libro B} + \text{Libro C} + \text{Libro O1} + \text{Cap. A1} + \text{Cap. A2} + \text{Cap. B} + \text{Cap. C} + \text{Cap. O1} + \text{Pat. Inv. A1} + \text{Pat. Inv. A2} + \text{Pat. Inv. A3} + \text{Pat. Inv. A4} + \text{Pat. Inv. O} + \text{Pat. Mod. Util. A} + \text{Pat. Mod. Util. O} + \text{Reg. A} + \text{Reg. O} + \text{Empresa A} + \text{NO. Pat. A} + \text{Norma.}$

Índice de producción de Nuevo Conocimiento Tipo A (NCA): Este indicador está conformado por un subconjunto de productos que tienen mayor

importancia y relevancia para la ciencia, y que se han publicado en revistas y editoriales reconocidas e indexadas de mayor impacto a nivel científico. Para este indicador los productos toman el mayor valor asignado, tal es el caso de los artículos en revistas tipo A que toman el mayor valor 1,0, libros 3,0, capítulos de libros 0,6, productos patentados 4,0, productos registrados 1,0, empresas spin off 1,0 y normas 1,0.

$\text{INCA} = \text{Art. A} + \text{Libro A1} + \text{Libro A2} + \text{Cap. A1} + \text{Cap. A2} + \text{Pat. Inv. A1} + \text{Pat. Inv. A2} + \text{Pat. Inv. A3} + \text{Pat. Inv. A4} + \text{Pat. Mod. Util. A} + \text{Reg. A} + \text{Empresa A} + \text{No. Pat. A} + \text{Norma.}$

Cabe destacar que los Índice de producción de Nuevo Conocimiento (NC) y el Índice de producción de Nuevo Conocimiento Tipo A (NCA) se diferencian porque, el primero agrupa todos los artículos, libros, capítulos, patentes, modelos de utilidad, creación de spin-off y normas que evidencien su existencia y calidad medido en diferentes niveles de impacto; el segundo índice, agrupa los mismos productos pero que cuentan con criterios de calidad de alto impacto a nivel nacional e internacional y clasificados por Colciencias como productos A y A1. Estos productos para considerarse A y A1 en el caso de los artículos deben estar publicados en revistas especializadas que se encuentran clasificadas en Publindex, Latindex, Scopus, ISI; los libros y capítulos de libros deben ser publicados por editoriales de alto impacto reconocidas por Colciencias y Book Citation Index; los software y patentes (Inventión y modelos de utilidad), deben estar registrados en las instituciones de registro correspondientes o de patentamiento formalmente constituidas; las empresas Spin-off deben tener registro tributario en Colombia, presentar estados financieros y al menos un producto en el mercado; y las normas, que corresponden a una regulación, reglamento o legislación, en este caso la entidad que emitió la norma debe certificar la participación del grupo en su construcción.

Índice de producción de Formación (F): Este indicador se refiere a los productos de formación del grupo, tales como tesis doctoral, tesis de maestría, trabajos de grado, programas académicos de doctorado, programa académico de maestría, cursos de doctorado y cursos de maestría, que se ponderan según sus características y peso entre 1,6 y 0,1.

$\text{IF} = \text{Tes. Doc. A} + \text{Tes. Doc. A2} + \text{Tes. Doc. A3} + \text{Tes. Doc. O} + \text{Tes. Maes. A} + \text{Tes. Maes. O} + \text{Trab. Gra. A} + \text{Trab. Gra. O} + \text{Prog. Aca. Doc.} + \text{Prog. Acad. Maes.} + \text{Curso Doc.} + \text{Curso de Maes.}$

Para validar los productos de formación, el miembro del grupo de investigación debe entregar copia del acta de grado, del diploma o del registro académico del estudiante en el que conste su participación como Director. La clasificación del producto en A, A2, A3 y O depende de la nota y el reconocimiento obtenido (Aprobado, Cum Laude, Summa Cum Laude). Así mismo, este índice tiene en cuenta los programas de doctorado y maestría y cada uno de los cursos creados por miembros del grupo de investigación.

Índice de producción de Divulgación (D): Se refiere a todos los productos de servicios y divulgación del grupo, como son: Servicios técnicos, consultorías, cursos de extensión, textos, literatura de circulación restringida, productos de divulgación. Las características y peso de los productos oscilan entre 1,0 y 0,3.

ID = Ser. Técnicos + Consult. + Curs. Ext. + Texto + Prod. Divulg. + Liter. Restrig.

Para validar los productos de divulgación se deben adjuntar el contrato del servicio de consultoría prestado, y certificados de calidad emitidos por la institución que utilizó el servicio, en los que conste el nombre del investigador, el grupo de investigación al que pertenece y la calidad de la consultoría.

El peso asignado a cada producto en este modelo es proporcional a la participación de los miembros del grupo. Es decir, cuando un producto tiene un único autor, sólo podrá estar relacionado a un grupo de investigación, cuando un producto es resultado de colaboración entre miembros de varios grupos, el peso total del producto se distribuye de manera proporcional teniendo en cuenta el número de autores del grupo. Después de obtener el peso de cada producto, se agrupan y se suman los pesos por tipología para obtener los indicadores de Nuevo Conocimiento (NC), Nuevo conocimiento tipo A (NCA), Formación (F) y Divulgación (D).

Para medir las variables independientes del modelo propuesto se diseñó y aplicó un **cuestionario** que se divide en 4 grandes bloques: datos generales, cultura organizacional, gestión del conocimiento y capital tecnológico; los enunciados de la encuesta relacionadas con las variables independientes se plantearon como afirmaciones y las respuestas en escala de estimación Likert (ordinal) con valores desde 1 (totalmente en desacuerdo), 2 (bastante en desacuerdo), 3 (neutral), 4 (bastante de acuerdo), 5 (totalmente de acuerdo).

Las variables de cultura de la organización, las afirmaciones propuestas y la escala Likert se

muestran en la tabla II; estas variables fueron propuestas basados en la taxonomía de valores y prácticas expuestas por diferentes autores en la literatura científica y agrupadas en 5 dimensiones: participativa, profesional, motivacional, trabajo en equipo, y emprendedora (Leidner y Kayworth, 2006; Pümpin y García, 1988; Moncaleano, 2002; Rueda, 2005).

Las variables de gestión del conocimiento se plantearon adoptando el modelo SECI (socialización, externalización, combinación e internalización), a partir del cual se plantearon las afirmaciones y la escala Likert, como se muestra en la tabla III (Bueno y otros, 2003; Jaime y otros, 2005; Gaviria y otros, 2007).

Las variables para el capital tecnológico se plantearon a partir del modelo Intellectus² (Bueno y otros, 2003), utilizado como referente para muchas investigaciones en otros estudios sobre organismos y centros de investigación en España, y que se muestran en la tabla IV (Bueno y otros, 2003; Meroño, 2005).

Para la variable producción científica se tomó como medida la categoría asignada por Colciencias (2010) de acuerdo a los resultados de investigación en un período de observación, como se explicó anteriormente y que se resume en la tabla V.

Para el análisis de los datos se utilizó el programa estadístico SPSS versión 18 y las técnicas estadísticas que se emplearon en este estudio se relacionan en la tabla VI.

La encuesta fue aplicada a los directores de grupos de investigación a través de Google Docs, obteniendo como respuesta 178 cuestionarios en formato online, y en otros casos dada la ubicación geográfica de algunas universidades se hizo la aplicación del cuestionario de forma personal obteniendo una respuesta de 45 cuestionarios, para un total de 223 encuestas.

Las variables finales del modelo se obtienen a partir del análisis factorial, técnica multivariante que parte del análisis de datos para proceder a realizar una agrupación de las variables similares en nuevas variables denominadas factores (Rueda, 2012). Esta reducción se requiere porque puede existir un exceso de variables que dificulta el análisis de la información, o porque presenta la misma información de manera redundante (Aldàs-Manzano, 2005). Las variables analizadas para la investigación presentaron cargas factoriales superiores a 0,70, un Alpha de Cronbach superior a 0,60, y la varianza explicada superior a 60%.

Tabla II. Variables e indicadores de Cultura de la Organización

CULTURA ORGANIZACIONAL Escala Likert (1-5)	
Variable e indicador	Afirmación
CULTURA PARTICIPATIVA: Orientada a la participación de los docentes investigadores en la planeación estratégica, la generación de ideas, facultad para tomar decisiones, comunicación y autonomía en el trabajo.	
Compromiso directivo	Los directivos de la universidad están comprometidos con los valores y prácticas orientados a la investigación.
Procesos de comunicación	Los procesos de comunicación en la universidad son asertivos y permiten el flujo de información.
Estilo de Dirección	El estilo de dirección en la universidad es participativo, flexible e integrador.
Toma de decisiones	La Dirección de Investigaciones en la universidad tiene autonomía en la toma de decisiones.
CULTURA PROFESIONAL: Orientada a la preparación y formación de los docentes investigadores, y la utilización de nuevas tecnologías en el ejercicio de la docencia y la investigación.	
Organización de Eventos científicos	La universidad promueve y apoya la realización de eventos científicos para la divulgación de la investigación.
Formación en programas de posgrado	La universidad apoya a los docentes investigadores con descarga de tiempo para la formación en programas de posgrado.
	La universidad brinda apoyo económico para la formación en programas de posgrado.
Asistencia a eventos científicos	La universidad brinda apoyo económico para asistir a eventos nacionales de investigación.
	La universidad brinda apoyo económico para asistir a eventos internacionales de investigación.
Formación en nuevas tecnologías	La universidad capacita a los docentes en nuevas tecnologías para el ejercicio de la investigación.
CULTURA MOTIVADORA: Se preocupa por el bienestar de las personas, posibilidades de desarrollo de los empleados, calidad y confianza en las relaciones con los directivos, reconocimiento y credibilidad en el trabajo realizado.	
Crecimiento profesional y personal	La universidad aporta a su crecimiento profesional y personal.
Incentivos económicos	La universidad ofrece incentivos económicos a los docentes investigadores por resultados anuales de producción científica.
Reconocimientos	La universidad realiza reconocimientos públicos a los docentes investigadores por logros anuales de producción científica.
CULTURA TRABAJO EN EQUIPO: Orientada al desarrollo de actividades en equipo interdisciplinarios dentro de la organización, que comparten conocimientos y experiencias de trabajo, participan en la toma de decisiones conjuntas, y generan valores como la colaboración y la solidaridad.	
Equipos multidisciplinares	La universidad promueve proyectos de investigación con equipos de trabajo multidisciplinares.
Nuevos grupos de investigación	La universidad promueve la creación de nuevos grupos y líneas de investigación.
Alianza entre grupos de investigación	La universidad fomenta las alianzas para la realización de proyectos entre grupos de investigación.
CULTURA EMPRENDEDORA: Orientada al desarrollo de nuevas experiencias y proyectos, mayor investigación, a asumir nuevos riesgos y tolerancia al fracaso, dispuestos al cambio y con pensamiento positivo frente a la innovación.	
Convenios interinstitucionales	La universidad promueve convenios con Universidad-Empresa-Estado para el desarrollo de proyectos de investigación.
Presupuesto para la investigación	La universidad asigna presupuesto suficiente para la investigación.
Políticas de investigación	La universidad promueve políticas generales que apoyan la investigación.
Nuevos proyectos	La universidad apoya y financia la ejecución de nuevos proyectos de investigación.

Fuente: Elaborado a partir de Pümpin y García, 1988; Moncaleano, 2002; Leidner y Kayworth, 2006; Rueda, 2005.

Tabla III. Variables e indicadores de Gestión del conocimiento

GESTIÓN DEL CONOCIMIENTO Escala Likert (1-5)	
Variable e indicador	Afirmación
SOCIALIZACIÓN	
Resultados del grupo	Los miembros del grupo socializan los resultados de investigación dentro del grupo.
Propuestas de investigación	Los miembros del grupo socializan las propuestas de investigación dentro del grupo.
Direccionamiento del grupo	Los miembros del grupo se reúnen para diseñar y direccionar estrategias que favorezcan el aprendizaje y crecimiento del grupo.
EXTERNALIZACIÓN	
Experiencias y resultados compartidos	El grupo de investigación comparte sus resultados y experiencias con otros grupos.
	El grupo de investigación registra y documenta los proyectos, resultados y experiencias investigativas.
Participación en redes científicas	El grupo de investigación participa en redes y sociedades científicas.
COMBINACIÓN	
Publicación de resultados de investigación	El grupo de investigación realiza la publicación de libros con el desarrollo de metodologías y resultados de investigación.
	El grupo de investigación realiza la publicación de artículos científicos.
Sistematización en Plataformas Tecnológicas	El grupo de investigación sistematiza los productos, resultados de proyectos de investigación en la plataforma virtual ScienTI-COL de Colciencias.
	El grupo de investigación sistematiza los productos, resultados de proyectos de investigación en la plataforma virtual propia de la universidad.
Registro de productos	El grupo de investigación registra ante la Dirección Nacional de Derechos de Autor el software desarrollado por el grupo.
	El grupo de investigación registra ante el organismo correspondiente las patentes desarrolladas por el grupo.
INTERNALIZACIÓN	
Uso de metodologías en el grupo	El grupo de investigación utiliza metodologías propias para el desarrollo de proyectos de investigación.
	El grupo de investigación utiliza métodos y experiencias para el ejercicio de la investigación, desarrollados por otros grupos.
Uso de resultados de investigación	El grupo de investigación tiene en cuenta los resultados de otros grupos de investigación para el desarrollo de sus propios proyectos.

Fuente: Elaborado a partir de Takeuchi, 2006; Gaviria y otros, 2007.

3.2 Caracterización de la muestra

De acuerdo a la muestra analizada, en el estudio participaron 73 universidades, con un total de 223 grupos de investigación que corresponden a 22 departamentos de Colombia, con un mayor porcentaje de participación de Santander con el 26,5%, Bogotá D.C con el 14,8% y Cundinamarca con el 11,2%. El 56,31% de los grupos de la muestra pertenecen a universidades privadas y el 43,69% pertenecen a Universidades públicas.

La tabla VII muestra el número de grupos que participaron en la muestra y el porcentaje de participación en cada una de las categorías definidas por Colciencias (2010).

Otro aspecto a destacar es el tamaño de los grupos de investigación, en la tabla VIII se muestra que los grupos con mayor producción científica tienen más personal de investigación y los grupos con menor producción científica cuentan con menos personal investigador. Además, los grupos con mayor producción científica tienen más de 10 años de creación y los grupos con menor producción científica están entre 1 y 10 años de creación.

A partir de la información registrada en la plataforma ScienTI-Col de Colciencias (GrupLAC) sobre los proyectos y productos realizados por los grupos de investigación de la muestra, en la tabla IX se encuentra un resumen del total de artículos producidos en cada categoría, libros, capítulos de libros, ponencias, software desarrollado y software registrado.

Tabla IV. Variables e indicadores de Capital Tecnológico

CAPITAL TECNOLÓGICO Escala Likert (1-5)	
Variable	Afirmación
INVERSIÓN I+D	
Personal de investigación	El grupo de investigación cuenta con personal suficiente para el desarrollo de proyectos de investigación.
Inversión en investigación	El grupo de investigación cuenta con suficiente presupuesto para la ejecución de proyectos de investigación.
DOTACIÓN TECNOLÓGICA	
Equipos y software	El grupo de investigación cuenta con software y equipos suficientes y adecuados para el desarrollo de la investigación.
Laboratorios	El grupo de investigación cuenta con laboratorios suficientes y adecuados para el desarrollo de la investigación.
Recursos bibliográficos	El grupo de investigación tiene disponibilidad de recursos bibliográficos de impacto científico.
HERRAMIENTAS TIC	
Bases de datos	Las bases de datos (académicas, científicas, empresariales) a las que tiene acceso la universidad, son suficientes y de calidad para el desarrollo de proyectos de investigación del grupo.
Herramientas colaborativas	La universidad promueve y provee a los grupos, herramientas tecnológicas para la interacción virtual con la comunidad científica.
HERRAMIENTAS TIC (Dicotómica: Si - No)	
Sistemas de Gestión del conocimiento	La universidad cuenta con una plataforma tecnológica en la que se registra el conocimiento que se gestiona desde el grupo.
	El grupo de investigación cuenta con una plataforma tecnológica en la que registra las actividades internas.
HERRAMIENTAS TIC (Pregunta de selección múltiple)	
Herramientas colaborativas	¿Qué herramientas colaborativas utilizan los miembros del grupo de investigación para el quehacer científico?

Fuente: Elaboración a partir de Bueno y otros, 2003; Meroño, 2005.

Tabla V. Medición de producción científica

PRODUCCION CIENTIFICA		
Variable SPSS	Descripción	Escala
CLAS_GRUP_INVES	<p>Clasificación asignada por Colciencias (2010) aplicando la siguiente formula:</p> $\text{ScientiCol} = 5 * (\text{NC}) + 3,5 * (\text{NCA}) + 1 * (\text{F}) + 0,5 * (\text{D})$ <p>INC= Art. A + Art. B + Art. C + Art. O + Libro A1 + Libro A2 + Libro B + Libro C + Libro O1 + Cap. A1 + Cap. A2 + Cap. B + Cap. C + Cap. O1 + Pat. Inv. A1 + Pat. Inv. A2 + Pat. Inv. A3 + Pat. Inv. A4 + Pat. Inv. O + Pat. Mod. Util. A + Pat. Mod. Util. O + Reg. A + Reg. O + Empresa A + NO. Pat. A + Norma.</p> <p>INCA= Art. A + Libro A1 + Libro A2 + Cap. A1 + Cap. A2 + Pat. Inv. A1 + Pat. Inv. A2 + Pat. Inv. A3 + Pat. Inv. A4 + Pat. Mod. Util. A + Reg. A + Empresa A + No. Pat. A + Norma.</p> <p>IF= Tes. Doc. A + Tes. Doc. A2 + Tes. Doc. A3 + Tes. Doc. O + Tes. Maes. A + Tes. Maes. O + Trab. Gra. A + Trab. Gra. O + Prog. Aca. Doc. + Prog. Acad. Maes. + Curso Doc. + Curso de Maes.</p> <p>ID = Ser. Técnicos + Consult. + Curs. Ext. + Texto + Prod. Divulg. + Liter. Restrig.</p>	Ordinal A1, A, B,C,D

Tabla VI. Técnicas estadísticas de la investigación

Análisis	Técnica
Caracterización de los grupos de investigación que participan en el estudio	Estadística Descriptiva
Agrupación de variables	Análisis factorial con método de rotación varimax
Análisis de correlación y comprobación de hipótesis	Análisis de correlaciones Análisis de regresión múltiple Modelo Lineal ANOVA Análisis de Caminos Análisis Estructural

Tabla VII. Grupos de la muestra de acuerdo a la clasificación de Colciencias

CLASIFICACION DE COLCIENCIAS	No. Grupos	% Participación
A1	31	13,90%
A	18	8,10%
B	42	18,80%
C	57	25,60%
D	75	33,60%
TOTAL	223	100,00%

Tabla VIII. Número de investigadores por Grupo

CLASIFICACION	NUMERO DE INVESTIGADORES					Total Grupos por Categoría
	1 a 5	6 a 10	11 a 15	16 a 20	21 o más	
A1	3	6	6	9	7	31
	8,30%	8,60%	12,50%	25,00%	20,58%	
A	1	5	3	4	5	18
	2,80%	7,20%	6,25%	11,10%	14,70%	
B	7	16	7	6	6	42
	19,40%	23,18%	14,58%	16,16%	17,64%	
C	6	18	11	12	10	57
	16,70%	26,08%	22,91%	33,33%	29,41%	
D	19	24	21	5	6	75
	52,80%	34,78%	43,75%	13,88%	17,64%	
Total Grupos por tamaño (223)	36	69	48	36	34	223
	16,14%	30,9%	21,5%	16,14%	15,24%	

Fuente: Elaborado a partir de la plataforma ScientI Colciencias, 2010

El tiempo empleado para las labores de investigación se encuentra registrado y actualizado en el curriculum de cada investigador (CvLac) en la plataforma de Colciencias. En la tabla X se presenta el promedio de investigadores por grupo y el número de horas semanales dedicadas a la investigación.

Así mismo, dentro de la variable capital tecnológico se tuvo en cuenta el uso de las tecnologías de la información y las comunicaciones (TIC) en los grupos de investigación, específicamente he-

rramientas de apoyo para las actividades de investigación. Al respecto, se encontró que el 11,6% afirma que siempre hace uso de las TIC para las labores de investigación, el 35,3% las utiliza bastante; sin embargo, hay un 25,6% que poco las utiliza y un 2,3% que no las utiliza. De acuerdo a una lista de herramientas, los grupos respondieron que las que más utilizaban son: las bases de datos, el correo electrónico y los buscadores; y las menos utilizadas, los wiki, marcadores y blog.

Tabla IX. Número de productos por grupo de la muestra (Ventana de observación 2006-2010)

Clasificación Colciencias	No. Grupos	No. Artículos	No. PONENCIAS	No. LIBROS	No. CAP. LIBROS	SD**
A1	31	4150	3188	434	563	72
	PPG*	134	103	14	18	2
A	18	1368	831	256	346	53
	PPG*	76	46	14	19	3
B	42	2235	1956	418	404	108
	PPG*	53	47	10	10	3
C	57	1915	1497	255	330	128
	PPG*	34	26	4	6	2
D	75	1034	874	174	101	240
	PPG*	14	12	2	1	3

*PPG: promedio de productos por grupo

**SD: Software desarrollado

Fuente: Elaborado a partir de la plataforma ScientI Colciencias, 2010.

Tabla X. Tiempo dedicado a la investigación

Clasificación	No. Grupos	No. Investigadores	Total Horas semana	Promedio hora/grupo	Promedio hora/ Investigador	Promedio Invest/grupo
A1	31	494	6378	205,74	12,91	16
A	18	356	3439	191,06	9,66	20
B	42	766	6123	145,79	7,99	18
C	57	783	5965	104,65	7,62	14
D	75	813	4318	57,77	5,31	10

Fuente: Elaborado a partir de la plataforma ScientI Colciencias, 2010.

4. RESULTADOS

Para la comprobación de las hipótesis, se utilizó la técnica estadística de Regresión lineal múltiple. Las hipótesis H1, H2 y H3 se comprueban con un R^2 corregida=15,5%, DW=1,843, F=11,199 y una significancia 0,000. Los coeficientes Beta permiten identificar que las variables que influyen directamente sobre la producción científica con un nivel de significancia * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$ son la cultura motivadora, el proceso de externalización, la dotación tecnológica y el tiempo de investigación.

Continuando con la comprobación de las hipótesis H4, H5 y H6, mediante el análisis de caminos se pretende probar relaciones indirectas de las demás variables con la producción científica utilizando la regresión múltiple. Los resultados obtenidos se muestran en la tabla XI.

Basados en los resultados y las ecuaciones obtenidas en el análisis de regresión y caminos, en

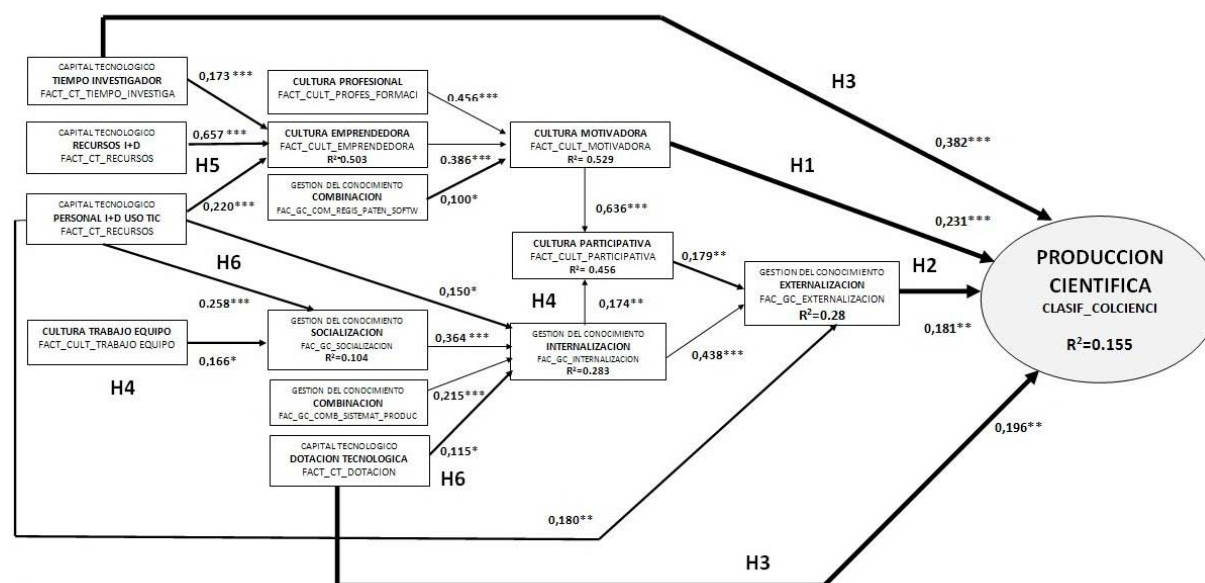
la figura 2 se presenta el modelo final con todas las relaciones encontradas entre las variables independientes Cultura Organizacional, Gestión del Conocimiento, Capital Tecnológico; y la variable dependiente Producción Científica.

El análisis de regresión comprueba las hipótesis H1, H2, H3. La cultura de la organización a través de la dimensión motivadora ($\beta=0,231$ ***), la gestión del conocimiento por medio del proceso de externalización ($\beta=0,181$ ***) y los elementos tiempo de investigación ($\beta=0,382$ ***) y dotación tecnológica ($\beta=0,196$ **) correspondientes al capital tecnológico, influyen de forma directa positiva y significativamente (* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$) sobre la producción científica ($R^2=0,155$).

Para la comprobación de las hipótesis H4, H5 y H6, se tomaron como variables dependientes cada una de las variables independientes. Este tipo de análisis, permitió encontrar relaciones directas positivas y sig-

Tabla XI. Resumen del Modelo final para la comprobación de las hipótesis

V. dependiente	V. Independientes	β	R ²	DW	F	Sig.
Producción Científica	Cultura motivadora Externalización Dotación Tecnológica Tiempo Investigación	0,231*** 0,181** 0,196** 0,382***	0,155	1,845	11,199	,000
Externalización	Cultura Participativa Internalización Personal I+D_ uso TIC	0,179** 0,438*** 0,180**	0,280	2,146	29,771	,000
Cultura Participativa	Cultura Motivadora Internalización	0,636*** 0,174**	0,456	1,815	92,162	,000
Cultura Motivadora	Cultura Profesional Cultura Emprendedora Combinación	0,456*** 0,386*** 0,100**	0,529	1,889	84,047	,000
Cultura Emprendedora	Recursos I+D Personal I+D_ uso TIC Tiempo Investigador	0,657*** 0,220*** 0,173***	0,503	1,976	75,972	,000
Internalización	Combinación Dotación tecnológica Personal I+D_ uso TIC Socialización	0,215*** 0,115* 0,150** 0,364***	0,283	1,764	22,930	,000
Socialización	Personal I+D_ uso TIC Cultura de equipo	0,258*** 0,166*	0,104	2,143	13,838	,000

Figura 2.

nificativas (* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$) entre otras variables del modelo y que por lo tanto, influyen de forma indirecta en la producción científica.

En este sentido, se puede decir que la cultura participativa ($\beta = 0,179^{**}$) referente al compromiso de los directivos por los valores y prácticas orientados a la investigación, estilo de dirección participativo y

flexible, autonomía en la toma de decisiones y los procesos de comunicación asertiva; el proceso de internalización ($\beta = 0,438^{***}$) que contempla el registro de metodologías propias, utilización y apropiación de métodos y resultados para la investigación; y el personal I+D que usa las TIC ($\beta = 0,180^{**}$); influyen positiva y significativamente sobre el proceso de externalización ($R^2 = 0,280$).

De igual modo, se encontró que la cultura motivadora ($\beta=0,636^{***}$) y el proceso de internalización ($\beta=0,174^{**}$) influyen sobre la cultura participativa ($R^2=0,456$). Así mismo, la cultura profesional orientada a la formación ($\beta=0,456^{***}$), la cultura emprendedora ($\beta=0,386^{***}$) y el proceso de combinación de desarrollo de patentes y registro de software ($\beta=0,100^{**}$) influyen positiva y significativamente sobre la cultura motivadora ($R^2=0,529$).

Otra relación encontrada, es la influencia del proceso de socialización ($\beta=0,364^{***}$), el proceso de combinación que abarca la sistematización de los productos de investigación ($\beta=0,215^{***}$), la dotación tecnológica ($\beta=0,115^*$), y el personal I+D que usa las TIC ($\beta=0,150^{**}$), influyen sobre el proceso de internalización ($R^2=0,256$). Teniendo en cuenta los anteriores resultados, se comprueba la Hipótesis H4: la cultura de la organización y la gestión del conocimiento se relacionan de forma positiva. Se destacan la cultura participativa y el proceso de internalización que influyen directamente sobre el proceso de externalización, que a su vez contribuye directamente sobre la producción científica.

Respecto a la H5, se encuentra evidencia positiva y significativa entre la cultura de la organización y el capital tecnológico. Los elementos del capital tecnológico, recursos I+D ($\beta=0,657^{***}$) referentes a gastos I+D, personal I+D, recursos bibliográficos y acceso a bases de datos científicas; el personal I+D que usa las TIC en la investigación ($\beta=0,220^{***}$), y el tiempo para la investigación por número de investigadores ($\beta=0,173^{***}$) se relacionan de forma positiva y significativa con la cultura emprendedora de las universidades referente a nuevos proyectos de investigación, asignación de presupuesto suficiente para la investigación y alianzas con empresas, universidades y estado ($R^2=0,503$).

La comprobación de la hipótesis H6 que plantea la relación positiva entre el capital tecnológico y la gestión del conocimiento. Se destaca la relación directa positiva y significativa existente entre el personal I+D y el uso de las herramientas colaborativas TIC con respecto a los procesos de socialización ($\beta=0,258^{***}$), internalización ($\beta=0,150^{**}$) y externalización ($\beta=0,180^{**}$); así mismo, la relación entre dotación tecnológica ($\beta=0,115^*$) y los procesos de internalización.

Finalmente, la confirmación de este modelo se hizo a través de análisis de ecuaciones estructurales que determinaron un índice NC ó Chi-normada (χ^2/gl) de 1,662 que se encuentra entre valores 1 y 2 y una significancia de 0,022 ($0,01 \leq p \leq 0,05$) para un buen ajuste del modelo. El valor RMSEA (*Root Mean Squared Error of Approximation*) es de 0,055 presentando un valor menor de $\leq 0,08$ que

se ajusta a los parámetros (Browne and Cudeck, 1992). Los valores GFI (Goodness of Fit Index) de 0,962, AGFI (Adjusted Goodness of Fit Index) de 0,929, NFI (Normed Fit Index) de 0,933, RFI (Relative Fit Index) de 0,90 y CFI (Comparative Fit Index) de 0,972, siendo todos mayores de 0,9 (Escrig-Tena y Bou-Llugar, 2005).

5. CONCLUSIONES

Los resultados obtenidos demuestran que la cultura de la organización a través de la dimensión motivadora (mayores incentivos y reconocimientos por los resultados obtenidos); la gestión del conocimiento a través del proceso de externalización (participación en sociedades científicas, compartición de conocimiento con otros grupos y centros de investigación, publicación de resultados de investigación); y el capital tecnológico a través de la dotación tecnológica (laboratorios, equipos de cómputo y software suficientes para la investigación) y el tiempo para la investigación (horas de investigación por número de investigadores); influyen directamente sobre los resultados obtenidos en producción científica por parte de los grupos de investigación.

De igual modo, se encuentra una fuerte relación entre las variables independientes cultura de la organización y la gestión del conocimiento, destacándose la relación entre la cultura participativa y el proceso de internalización que influyen directamente sobre el proceso de externalización, y que a su vez contribuyen directamente sobre la producción científica; y una positiva y significativa relación entre la cultura de la organización y el capital tecnológico. Los elementos del capital tecnológico recursos I+D (gastos I+D, personal I+D, recursos bibliográficos y acceso a bases de datos científicas), el uso de herramientas colaborativas de investigación (personal I+D y uso de las TIC) y el tiempo para la investigación (horas de investigación por número de investigadores) se relacionan de forma positiva y significativa con la cultura emprendedora de las universidades referente a nuevos proyectos de investigación, asignación de presupuesto suficiente para la investigación y alianzas con empresas, universidades y estado.

Los resultados encontrados en esta investigación se encuentran soportados en estudios realizados en diferentes países y contextos. Harvey y otros, (2002) coinciden en destacar la importancia de la cultura organizacional en los resultados de los grupos de investigación, especialmente lo referente a sistemas adecuados de motivación y retención del talento humano. Avital y Collopy (2001) también son un referente, ya que en su in-

vestigación concluyeron que existen factores institucionales asociados a la producción científica, tales como el tiempo de dedicación para la investigación, disponibilidad de recursos, y visibilidad de los resultados de investigación, entre otros. De igual forma, Bozeman y Corley (2004) afirman que las actividades de colaboración propician la gestión de nuevo conocimiento, disponibilidad de nuevos recursos, adopción de nuevas técnicas de investigación, movilidad e intercambio de investigadores. Gaviria y otros, (2007) en su estudio sobre gestión del conocimiento en los grupos de investigación en Colombia, coinciden en que las prácticas, rutinas, los procesos de formación de docentes investigadores, sistemas de incentivos, y trabajo interdisciplinario, influyen en una mejor gestión del conocimiento y el crecimiento de la producción científica. Finalmente, Acevedo y otros, (2005) encontraron que para el buen ejer-

cicio de la investigación, se debe tener en cuenta el mejoramiento de la infraestructura física y tecnológica (laboratorios, centros de documentación, bases de datos, acceso a redes, etc.), el personal cualificado para la investigación, y el aumento de estímulos para los investigadores.

Por lo anterior, los grupos de investigación para obtener mejores resultados en producción científica requieren de una cultura organizacional que promueva políticas, valores y prácticas alineadas a los resultados de investigación; mayor asignación de recursos tecnológicos, tiempo para ejercer la investigación y personal I+D que contribuyan al desarrollo de la investigación y a una mejor gestión del conocimiento; de esta manera se podría obtener una mayor producción científica y una mejor clasificación para los grupos de acuerdo a la evaluación que realiza Colciencias en Colombia.

6. NOTAS

[1] El Modelo SECI planteado por Nonaka y Takeuchi (1995), define el proceso de conversión e interacción de conocimiento tácito y explícito en cuatro modos distintos: Socialización; conocimiento adquirido a través de la experiencia y compartida con otras personas de forma tácita; Externalización, conversión del conocimiento tácito a explícito, haciéndolo formal y fácil de compartir a muchas más personas; Combinación, conocimiento explícito que se distribuye a través de sistemas formales y más complejos basados en tecnologías de la información y la comunicación; Internalización, conversión del conocimiento ex-

plícito a tácito en el ejercicio de aprender haciendo, logrando la interiorización de documentos y metodologías, del saber al hacer; de esta manera el modelo es una espiral que lleva a que el conocimiento sea dinámico y continuo.

[2] El modelo Intellectus desarrollado por Bueno y otros (2003) establece cinco tipos de capital intelectual: capital humano, capital organizativo, capital tecnológico, capital de negocio, y capital relacional. El capital tecnológico, permite medir todo lo concerniente a esfuerzo I+D y dotación tecnológica en los centros y organismos de investigación.

7. REFERENCIAS

- Acevedo, M.; González, O.; Zamudio, L.; Abello, R.; Camacho, J.; Gutiérrez, M.; y otros. (2005). Un análisis de la transferencia y apropiación del conocimiento en la investigación de universidades colombianas. *Investigación y Desarrollo*, 13(1), 128-157.
- Aldàs-Manzano, J. (2005). *Análisis Multivariable aplicado*. Madrid: Thomson.
- Arbonés, A. (1999). Case studies as a tool for the externalization of tacit managerial Knowledge. *Paper WCRA*. En Gaviria y otros (2007), Gestión del conocimiento en los grupos de investigación de excelencia de la Universidad de Antioquia.
- Avital, M.; Collopy, F. (2001). *Assessing Research Performance: Implications for Selection and Motivation*. Case Western Reserve University, USA. Recuperado el 2010, de Sprouts: Working Papers on Information Systems: <http://sprouts.aisnet.org/1-14>
- Bermeo, H. (2007). Rendimiento y Colaboración Científica en la investigación académica. Estudio del Caso de los Grupos de Investigación de la Universidad Politécnica de Valencia. *Tesis Doctoral*. Valencia, España: Departamento de proyectos de ingeniería. Universidad Politécnica de Valencia.
- Bozeman, B.; Corley, E. (2004). Scientists' collaboration strategies: implications for scientific and technical human capital. *Research Policy*(33), 599-616. <http://dx.doi.org/10.1016/j.respol.2004.01.008>
- Browne, M.W.; Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, 21 (2), 230 - 258. <http://dx.doi.org/10.1177/0049124192021002005>
- Bueno, E.; Morcillo, P.; Rodríguez, J.; Luque, M. Á.; Cervera, M.; Camacho, C.; y otros. (2003). Gestión del Conocimiento en universidades y organismos públicos de investigación. D. G. *Investigación, C. d. Educación, & C. d.* Madrid, Edits. 1-60.

- Chang Lee, K.; Sangjae, L.; Won Kang, I. (2005). KMPi: measuring knowledge management performance. *Information & Management*(42), 469-482. <http://dx.doi.org/10.1016/j.im.2004.02.003>
- Cherchye, L.; Vanden Abeele, P. (2005). On research efficiency A micro-analysis of Dutch university research in Economics and Business Management. Belgium: Catholic University of Leuven, Campus Kortrijk and Faculty of Economics and Applied Economics. *Research Policy* 34 (2005) 495-516. <http://dx.doi.org/10.1016/j.respol.2005.03.005>
- Clark, B. (1997). *Las universidades modernas: Espacios de investigación y docencia*. México: Miguel Angel Porrúa - UNAM.
- Colciencias. (2008). *Modelo de Medición de Grupos de Investigación, Tecnológica o de Innovación*. Recuperado el 18 de Febrero de 2011, de <http://www.colciencias.gov.co/sites/default/files/upload/documents/2656.pdf>.
- Colciencias. (2010). Informe Ejecutivo: Convocatoria Nacional para la medición de Grupos de Investigación, Tecnológica o de Innovación. Recuperado el 18 de febrero de 2011, de <http://www.colciencias.gov.co>.
- Czerniewicz, L.; Brown, C. (2009). A study of the relationship between institutional policy, organisational culture and e-learning use in four South African universities. *Computers & Education* , 53, 121-131. <http://dx.doi.org/10.1016/j.compedu.2009.01.006>
- Escrig-Tena, A.B.; Bou-Lluisar, J.C. (2005). A model for evaluating organizational competencies: An application in the context of a quality management initiative. *Decision Sciences*, 36 (2), 221-257. <http://dx.doi.org/10.1111/j.1540-5414.2005.00072.x>
- Gaviria Velásquez, M.; Mejía Correa, A.; Henao Henao, D. (2007). Gestión del conocimiento en los grupos de investigación de excelencia de la Universidad de Antioquia. *Revista Interamericana de Bibliotecología, Jul.- Dic.*, 30(2) , 137-163.
- Harvey, J.; Pettigrew, A.; Ferlie, E. (2002). The Determinants of Research Group Performance: Towards Mode 2? *Journal of Management Studies*, vol. 39: pág. 747-774. <http://dx.doi.org/10.1111/1467-6486.00310>
- Heinze, T.; Shapirab, P.; Rogers, J.; Senk, M. (2009). Organizational and Institutional Influences on Creativity in Scientific Research. *Research Policy*, 610-623. <http://dx.doi.org/10.1016/j.respol.2009.01.014>
- Jaime, A.; Blanco, C. (2007). La gestión de conocimientos en entidades de conocimiento. El caso de los laboratorios académicos y de las empresas de base tecnológica en Europa. *Pensamiento y gestión*(22), 168-190.
- Jaime, A.; Gardoni, M.; Mosca, J.; Vinck, D. (2005). BASIC Lab: a software tool for supporting the production of knowledge in research organizations through the management of scientific concepts. *Journal of Knowledge Management*, 9(6), 53-66. <http://dx.doi.org/10.1108/13673270510629963>
- Leidner, D.; Kayworth, T. (2006). A review of culture in information systems research: "Toward a theory of information technology culture conflict". *MIS Quarterly*, Vol 30, No. 2, pp 357-399.
- Lopez-Nicolas, C.; Meroño, A. (2009). The impact of Organizational culture on the use of ICT for Knowledge management. *Eletron Markets* (219), 211-219. <http://dx.doi.org/10.1007/s12525-009-0020-4>
- Meroño, A. (2005). Tecnologías de información y Gestión del conocimiento: Integración de un sistema. *Economía Industrial*(357), 107-116.
- Miquel, S. (1997). *Investigación de mercados*. España: McGraw-Hill.
- Moncaleano, G. I. (2002). La tecnología de la información en organizaciones cooperativas: influencia sobre el aprendizaje, la creación de valor, y cultura. *Tesis doctoral*. España: Universidad Politécnica de Valencia.
- Nonaka, I.; Takeuchi, H. (1995). *The knowledge creating company. How Japanese companies create the dynamics of innovation*. New York: Oxford University Press.
- Plataforma ScienTI-Colombia. (s.f.). Recuperado en marzo/noviembre de 2011, de <http://www.colciencias.gov.co/scienti>
- Pümpin, C.; García, S. (1988). *Cultura Empresarial*. Madrid: Díaz de Santos.
- Rivera, R.; Sampedro, J. L.; Dutrénit, G. (2009). How productive are academic researchers in agriculture-related sciences? The Mexican case. *Working paper series*.
- Rodríguez, M.; Paéz, D. Influencia de la cultura social y de la cultura organizacional en las universidades. *Boletín de Psicología*(97), 9-34.
- Rueda Barrios, G. E. (2012). Influencia de la cultura organizacional, la gestión del conocimiento y el capital tecnológico en la producción científica. Aplicación a grupos de investigación adscritos a Universidades en Colombia. *Tesis Doctoral*, Universidad Politécnica de Valencia, Organización de Empresas, Valencia. <http://dx.doi.org/10.4995/thesis/10251/15666>
- Rueda, G. (2005). Influencia de la Cultura de la empresa en los resultados de innovación. (U. P. Catalunya, Ed.) *3er Congreso de Soporte del conocimiento con la tecnología SOCOTE*, 117 - 129.
- Tomàs, M.; Rodríguez, D. (2009). Conocer la cultura de la universidad contemporánea: el CICOU. *Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI)*.
- Van der Meulen, B. (2002). Europeanization of Research and the Role of Universities: an Organizational-Cultural Perspective, *Innovation: The European Journal of Social Science Research*, 15 (4): 341-356. <http://dx.doi.org/10.1080/1351161022000042570>



ESTUDIOS / RESEARCH STUDIES

Ranking de las bibliotecas universitarias españolas en la gestión del personal

José Simón-Martín*, Clara Simón-Blas**, Alicia Arias-Coello***

*Instituto Universitario de Evaluación Sanitaria. Universidad Complutense de Madrid. Facultad de Medicina.

**Departamento de Estadística e Investigación Operativa. Universidad Rey Juan Carlos.

***Facultad de Ciencias de la Documentación. Universidad Complutense de Madrid.

Correo-e: aarias@ucm.es

Recibido: 04-02-2015; 2ª versión: 20-05-2015; Aceptado: 25-05-2015.

Cómo citar este artículo/Citation: Simón-Martín, J.; Simón-Blas, C.; Arias-Coello, A. (2016). Ranking de las bibliotecas universitarias españolas en la gestión del personal. *Revista Española de Documentación Científica*, 39(1): e119. doi: <http://dx.doi.org/10.3989/redc.2016.1.1259>

Resumen: Los objetivos de este trabajo son dos. En primer lugar, identificar aquellas bibliotecas universitarias públicas españolas que durante los años 2008 y 2012 han sido las mejores en el ámbito de la eficiencia en su gestión del personal y que, por tanto, pueden ser consideradas como referentes en la mejora de la eficiencia; en segundo término comparar las diferencias existentes en la organización del personal entre las bibliotecas más ineficientes y las más eficientes en dicha área. El cálculo de la eficiencia se ha realizado mediante la técnica de análisis envolvente de datos (Data Envelopment Analysis-DEA); así mismo, para el establecimiento de un ranking entre las bibliotecas eficientes se han utilizado los métodos de contar el número de unidades de referencia, el cálculo de la super-eficiencia y la evaluación cruzada. Las bibliotecas analizadas muestran en su conjunto una eficiencia global y técnica pura alta. De entre las bibliotecas eficientes, dos de ellas ocupan el primer lugar de la clasificación en los dos años estudiados. Estas dos bibliotecas disponen de una organización del personal diferente a las dos consideradas como más ineficientes.

Palabras clave: Análisis envolvente de datos (DEA); bibliotecas universitarias; eficiencia; super-eficiencia; eficiencia cruzada; personal bibliotecario; eficiencia del personal; ranking bibliotecario.

Ranking of Spanish university libraries according to their personnel management

Abstract: The aim of this work is twofold: first, to identify those Spanish public university libraries that have been the best at efficiently managing their personnel between 2008 and 2012, and can therefore be considered a reference in the area of efficiency improvement. Secondly, to look at the differences in the way personnel is organized between the most and the least efficient libraries. Efficiency was calculated using a Data Envelopment Analysis (DEA). In addition, the method used to rank the most efficient libraries consisted of counting the number of reference units, calculating the super-efficiency scores, and applying a cross-evaluation method. On the whole, the libraries analysed show a high global and pure technical efficiency. Of the most efficient libraries, there are two that are at the top of the ranking in each of the years studied. Personnel organization in these two libraries is different from that of the two least efficient centres in the study.

Keywords: Data Envelopment Analysis (DEA); university libraries; efficiency; super-efficiency; cross-efficiency; library personnel; personnel efficiency; library ranking.

Copyright: © 2016 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0.

1. INTRODUCCIÓN

Desde el año 2008 España vive una profunda crisis económica que produce en las universidades públicas una notable reducción en su financiación. Ante este hecho, las universidades vienen desarrollando programas para mejorar su eficiencia con el objetivo de conseguir el equilibrio presupuestario.

Esta situación de crisis financiera tiene también, naturalmente, repercusiones en las bibliotecas universitarias, las cuales han visto reducidas sus partidas presupuestarias para la adquisición de materiales bibliográficos, así como para la tasa de reposición del personal. En este escenario, las bibliotecas también deben adaptarse para mejorar, entre otras cosas, las políticas de compras, así como la eficiencia de sus recursos humanos cuyo coste viene a representar entre el 35 y el 40% del presupuesto bibliotecario (Jarrillo Calvarro, 2012). Con el fin de que los gestores puedan mejorar la eficiencia del personal, les sería muy útil conocer, por un lado, el nivel de eficiencia con el que su personal está llevando a cabo los procesos bibliotecarios y, por otro, identificar aquellas bibliotecas que sean referentes en esta área.

El análisis envolvente de datos (Data Envelopment Analysis, conocido por sus siglas DEA), ha sido una de las técnicas utilizada por numerosos investigadores para el cálculo de la eficiencia de las organizaciones. El DEA, introducido por Charnes y otros (1978), está basado en una metodología que mide la eficiencia relativa de un grupo de organizaciones, clasificando a éstas en eficientes e ineficientes. El criterio que utiliza para esta clasificación viene determinado por la localización de cada una de esas organizaciones con respecto a la frontera de producción eficiente. Esta frontera la forman las unidades que tienen mejores prácticas, en referencia a la relación entre los productos ya elaborados y los recursos empleados en su producción. Las unidades identificadas como ineficientes se pueden comparar con una o más unidades localizadas en la frontera, las cuales pueden considerarse como sus unidades de referencia o *benchmarks* para orientar sus esfuerzos de mejora (Bergendahl, 1998).

Dado que como resultado del empleo del DEA disponemos, generalmente, de más de una unidad eficiente, se impone el establecer un ranking entre ellas con el objeto de poder discriminar mejor aquellas bibliotecas que deberían ser utilizadas como modelos de referencia. En la literatura se han propuesto diversas metodologías para el establecimiento de rankings, y para una revisión general de los mismos se pueden consultar los trabajos de Adler y otros (2002), Angulo- Meza y Lins (2002), y Lu y Lo (2009).

Para la realización del presente trabajo se han utilizado los tres métodos más empleados en la literatura con el fin de conseguir una discriminación dentro del grupo de bibliotecas eficientes: estudio de la frecuencia con la que aparece un centro eficiente como referencia de los centros ineficientes, el cálculo de la super-eficiencia introducido por Andersen y Petersen (1993) y la evaluación cruzada (cross-evaluation), propuesto inicialmente por Sexton y otros (1986). Esta última técnica, aunque ha sido aplicada a numerosos sectores productivos, hasta el momento no ha sido utilizada en el ámbito bibliotecario.

La técnica del DEA ha sido empleada por numerosos investigadores para calcular la eficiencia de las bibliotecas universitarias (Chen, 1997; Mann, 1997; Kao y Lin, 1999; Kao y Liu, 2000; Shim, 2000 y 2003; Saunders, 2003; Kao y Lin, 2004; Reichmann, 2004; Reichmann y Sommersguter-Reichmann, 2006 y 2010; Simon-Blas y otros, 2007; Stancheva y Angelova, 2008; Liu y Chuang, 2009; Lai y otros, 2011; Owadie, 2011; Simón y otros, 2011; Noh, 2011; Hwang y otros, 2012; Carvalho y otros, 2012-2013; Lee y otros, 2013; Shahwan y Kaba, 2013). Sin embargo, de los anteriores trabajos solo tres hacen referencia a la aplicación del DEA en *benchmarking*. El primero y más completo es el realizado por Reichmann y Sommersguter-Reichmann (2006) en el que se analizan 118 bibliotecas universitarias y se establece una ordenación de las unidades eficientes mediante la medición del número de veces que éstas eran referentes de unidades ineficientes, el cálculo de la super-eficiencia y el cálculo de un índice desarrollado por Torgersen y otros, (1996). Posteriormente aparecieron otros dos artículos en los que únicamente se consideraba el número de veces que cada biblioteca eficiente era referente de una biblioteca ineficiente para establecer un ranking entre ellas (Lai y otros, 2011; Noh, 2011).

En este artículo se pretende calcular, en una muestra de 41 bibliotecas universitarias españolas, la eficiencia global, técnica y de escala del personal en relación con las actividades que realiza en la biblioteca, estableciendo un ranking entre aquellas bibliotecas identificadas como eficientes y comparando la forma en que están organizados los recursos humanos de las mismas teniendo en cuenta las más eficientes frente a las más ineficientes.

En el estudio se ha tomado el año 2008 como referencia para el análisis de la eficiencia, comparando estos resultados con los obtenidos en 2012, con el objetivo de verificar si la crisis económica ha tenido algún impacto en los valores de eficiencia.

2. METODOLOGÍA

2.1. Cálculo de la eficiencia

El nombre de DMU (decision making unit) fue utilizado por primera vez por Charnes y otros (1978), para describir las unidades que van a ser analizadas mediante la técnica DEA. En el presente artículo, cada biblioteca representa una única DMU.

La eficiencia de una organización se define como la razón de la suma ponderada de los productos producidos (outputs), con respecto a la suma ponderada de los recursos utilizados para su producción (inputs), siendo el propósito del DEA hacer que el valor de eficiencia para cada unidad en la muestra sea el máximo que pueda alcanzar. Para ello, se ajustan los pesos de la combinación de variables de inputs y de outputs, de acuerdo con el resto de las unidades de la muestra. Siguiendo a Charnes y otros (1978), dada la unidad (DMU_j), el objetivo es conseguir la máxima eficiencia de dicha unidad (E_j), tal y como expresa el modelo de la ecuación (1):

$$\text{Max } E_j = \frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \quad \forall j = 1, \dots, n \quad (1)$$

- donde m es el número de variables del input y s es el número de variables del output
- y_{rj} es el valor de la variable r del output en la DMU j -ésima para $r = 1 \dots s$;
- x_{ij} es el valor de la variable i del input en la DMU j -ésima para $i = 1 \dots m$;
- u_r es el peso de la variable r -ésima del output;
- v_i es el peso de la variable i -ésima del input; y
- n es el número de unidades de decisión (DMUs).

El objetivo, como hemos señalado anteriormente, es optimizar los pesos u y v para que el valor de eficiencia resultante para esa unidad de análisis sea el mayor posible.

Las restricciones a este modelo son las siguientes:

- Los valores de eficiencia están acotados entre 0 y 1 (ya que son números relativos) y, por tanto:

$$0 \leq \frac{\sum_r u_r y_{rj}}{\sum_i v_i x_{ij}} \leq 1 \quad \forall j = 1, \dots, n \quad (2)$$

- Los pesos que se utilicen deben ser valores no negativos y, por tanto:

$$u_r, v_i > 0 \quad \forall r = 1, \dots, s; i = 1, \dots, m \quad (3)$$

Resolviendo el problema de programación lineal para cada una de las DMU, se puede calcular el conjunto de ponderaciones de las variables del input y de las del output que permiten obtener el mayor índice de eficiencia.

Las DMUs con el valor máximo de eficiencia serán aquellas cuyo E_j alcanza el valor de 1 y se les denomina unidades eficientes, las cuales constituirán el conjunto de referencia para las unidades ineficientes.

Este modelo, al comparar los valores de las variables del input y del output de las unidades ineficientes con los valores de las unidades eficientes que les sirven de referencia, nos permite conocer sobre qué variables debemos actuar, así como la intensidad de dicha actuación, con el objetivo de que esas unidades ineficientes alcancen la eficiencia.

El modelo definido por las ecuaciones (1, 2 y 3) se denomina de escala con retorno constante CRS (**C**onstant **R**eturn **S**cale), esto significa que un incremento en el volumen de los inputs o variables de entrada va a producir un incremento proporcional del volumen de los outputs o variables de salida. Posteriormente, Banker y otros (1984), ampliaron dicho modelo con la finalidad de tener en cuenta los efectos de la escala de funcionamiento de las DMUs. A dicho modelo se le denomina de escala con retorno variable VRS (**V**ariable **R**eturn **S**cale), esto es, un incremento en los inputs producirá un incremento proporcionalmente mayor o menor de acuerdo con la escala de funcionamiento de dicha DMU. El valor de eficiencia calculado con el modelo VRS, que denotaremos por E_{VRS_j}, se denomina "eficiencia técnica pura", y la ratio entre ésta y el valor de eficiencia calculado con el modelo CCR (conocido como "eficiencia global" y que denotaremos por E_{CCR}), se denomina "eficiencia de escala" (ES):

$$ES = (E_{VRS} / E_{CCR})$$

Cuando el valor de la eficiencia de escala (ES) es igual a 1, significa que la DMU está operando en el tamaño de escala más productivo. Alternativamente, si el valor de (ES) es menor a 1, entonces la DMU puede estar operando en una escala de retorno decreciente, lo que significa que un incremento del volumen del input produce un menor incremento del output al que le correspondería si esta relación fuera proporcional. Sucedería lo contrario si la DMU estuviera funcionando en una escala con retorno creciente.

Otro elemento importante en el uso de esta metodología es el tipo de orientación que se va a utilizar en el cálculo de la eficiencia. En el modelo de orientación al input (recursos), queremos indicar que una unidad ineficiente puede transformarse en

eficiente mediante la reducción del volumen de sus recursos (input), al tiempo que mantiene el mismo nivel de volumen de sus productos (output). En el caso de una orientación al output, el objetivo sería el conseguir un incremento del volumen de productos (output), manteniendo el volumen de los recursos utilizados para su producción (input).

En este trabajo hemos utilizado una orientación al input al considerarla la más apropiada en el sector de la biblioteca, dado que los gestores tienen más posibilidades de actuar sobre los recursos de los que disponen, que de modificar el volumen de los servicios que producen.

2.2. Cálculo de la super-eficiencia

Andersen y Petersen (1993) desarrollaron un método para clasificar las unidades eficientes. Esta metodología permite que una unidad eficiente pueda alcanzar un valor superior a 1. Para ello se recalcula la distancia radial de cada unidad eficiente a la frontera de producción, construida por las restantes unidades una vez eliminada ésta de la muestra. Esto se consigue eliminando la restricción (2) del modelo de programación lineal expresado por las ecuaciones (1, 2, y 3).

2.3. Cálculo de la eficiencia cruzada (Cross-efficiency)

En este método inicialmente desarrollado por Sexton y otros (1986), cada unidad es evaluada con las ponderaciones más óptimas de las restantes para el cálculo de la eficiencia, construyéndose una matriz de eficiencia cruzada. Una vez obtenida esta matriz, cada unidad tiene, por un lado, su propia autoevaluación y, por otro, dispone de la evaluación que ha recibido utilizando las ponderaciones de las otras unidades de la muestra (evaluación por sus pares) y también los valores de eficiencia obtenidos por las unidades cuando son evaluadas con sus ponderaciones (evaluación de los pares). Para un desarrollo detallado de la metodología empleada véase el trabajo de Doyle y Green (1994).

En este estudio hemos utilizado como valor de eficiencia cruzada de una unidad, el del promedio obtenido de las valoraciones realizadas por sus pares. Por tanto, una unidad con un alto valor de eficiencia cruzada supone que no solo es buena por sí misma, sino que es considerada también buena por la mayoría de las otras unidades.

Posteriormente Doyle y Green (1994), desarrollaron un índice (índice de "maverick"), para medir la desviación entre la eficiencia de la unidad y el promedio del valor obtenido mediante la "evaluación por pares" en la matriz de eficiencia cruzada.

Una unidad con un índice de "maverick" elevado, es una unidad anómala dentro de su conjunto de referencia (Doyle y Green, 1994) y posiblemente sea un "falso positivo" (Talluri y Sarkis, 1997), que ha conseguido ser eficiente debido al uso de unas ponderaciones inapropiadas para las variables.

2.4. Organización de los recursos humanos bibliotecarios

Para conocer la manera en la que cada biblioteca tiene organizados sus recursos humanos, se ha procedido al análisis de su organigrama, bien a partir de la información que aparece en su página web, o bien analizando la organización del personal adscrito a la biblioteca según se describe en la relación de puestos de trabajo aprobada por cada universidad a la que corresponden cada una de ellas.

3. SELECCIÓN DE BIBLIOTECAS Y VARIABLES A UTILIZAR EN EL ANÁLISIS

3.1. Selección de la muestra de bibliotecas a analizar

Para la realización del presente trabajo se seleccionaron las 47 bibliotecas pertenecientes a las universidades públicas españolas de tipo presencial, que se redujeron a 41 por las razones que exponemos a continuación. Los datos utilizados en este estudio se recogieron del anuario estadístico publicado en la página web de REBIUN referidos a los años 2008 y 2012. La muestra quedó reducida en primer lugar a 46 bibliotecas dado que hubo que eliminar a la biblioteca de la Universidad Politécnica de Cartagena, ya que no presentaba datos en el año 2012.

Con el fin de mejorar la fiabilidad de los datos que aparecen en las estadísticas de REBIUN, se han utilizado tres métodos. El primero consistió en tener en cuenta todos los casos y seleccionar, para cada uno de los datos, los valores que habían tenido durante los últimos 5 años, con el fin de comprobar la no existencia de cifras erráticas en la serie. El segundo método se basó en la verificación de la bondad del dato de REBIUN, comparándolo con el que aparecía en la memoria de la biblioteca, aunque aquí hemos de aclarar que esta última aproximación sólo se pudo llevar a cabo en los casos en los que se disponía de una memoria anual de actividades. El tercer método utilizado se centró en la identificación de aquellas bibliotecas que son atípicas; esto es, bibliotecas que presentan datos extremos en comparación con el resto. La presencia de datos atípicos puede ser debida a diversos factores: resultado de errores en la medición, presencia de características inusuales con

respecto al resto de la muestra, o la no pertenencia al grupo de referencia según el cual la unidad está siendo evaluada (Johnson y Mc Ginnis, 2008). La importancia de detectar unidades extremas reside en que los estimadores del DEA son muy sensibles a la presencia de éstos, ya que el método emplea observaciones extremas para identificar un desempeño superior (Sexton y otros, 1986).

Todas las bibliotecas atípicas presentan la característica de que son de gran tamaño, siendo ineficientes en el modelo de DEA con escala de retorno constante (CRS) para ambos años. Sin embargo, cuando se utiliza el modelo de DEA con escala de retorno variable todas ellas son eficientes debido a que una, o varias de sus variables, alcanzan valores extremos.

Para la detección de unidades atípicas o extremas, se han propuesto diversos métodos en la literatura (Simar, 2003; Banker y Chang, 2006; Johnson y Mc Ginnis, 2008; Bellini, 2012). En este estudio hemos utilizado el método propuesto por Bellini (2012), que permite la identificación de unidades atípicas que se localizan dentro y fuera de la frontera de eficiencia. Para ello, se construye para cada output una función de distancia que considera los valores de súper-eficiencia de cada unidad de referencia y el valor del output de dicha unidad. Con el fin de determinar las unidades atípicas, en este artículo hemos considerado la elipse al 95% del boxplot bivariado, que resulta de considerar las distancias mínimas y máximas de los outputs a los valores de súper-eficiencia alcanzados por cada biblioteca (figura 1).

Con el empleo de este procedimiento se ha detectado que existen dos bibliotecas, tanto en 2008 como en 2012, que se localizan fuera de la elipse que representa la curva segmentaria del 99% (las bibliotecas de las universidades Complutense de Madrid y la de Valencia). En el año 2008 aparece otra biblioteca, la de la Universidad de Zaragoza, que está localizada fuera de la elipse que representa la curva segmentaria del 95%. Así mismo, en 2012, y en esa misma elipse, se localizaron las bibliotecas de las universidades de Barcelona y de Salamanca; por tanto, la muestra final de bibliotecas que se ha analizado para la realización de este trabajo, ha sido de 41, cifra que resulta de la eliminación de las anteriores bibliotecas por los motivos expuestos.

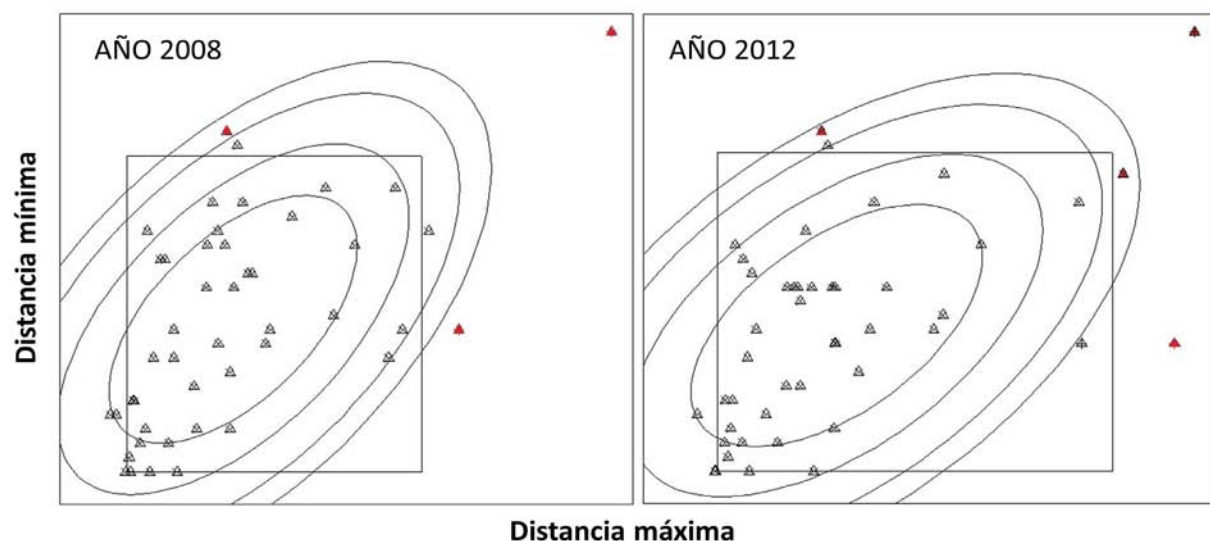
3.2. Selección de las variables

En la técnica DEA la identificación y selección de los inputs y outputs más adecuados es fundamental, y va a depender de su relevancia de acuerdo con los objetivos de la evaluación y de la factibilidad de su obtención. En este trabajo hemos analizado el conjunto de procesos en los que se precisa una participación relevante de los recursos humanos, y en los que las políticas acerca de la organización de los mismos puedan tener importancia en el cálculo de su eficiencia.

Los recursos humanos de una biblioteca universitaria se utilizan fundamentalmente en los dos siguientes macroprocesos:

- a) El de gestión de la colección, que tiene el objetivo de crear colecciones adecuadas a las necesidades

Figura 1. Boxplot de los outputs de los valores de súper-eficiencia alcanzados por cada biblioteca en 2008 y 2012. En rojo bibliotecas atípicas



de aprendizaje, docencia e investigación de la universidad, y que incluye tres grandes procesos: adquisiciones, técnico y mantenimiento de la colección. Como aproximación a las cargas de trabajo de estas actividades se han considerado:

- Actividades relacionadas con el mantenimiento de la colección para su uso inmediato. Como estimador del volumen de dichas actividades podemos considerar el volumen del fondo informatizado del que dispone la biblioteca, y que hemos denominado "ITEMS".
 - Actividades relacionadas con la identificación de necesidades de los usuarios, selección y adquisición de materiales. Como estimador del volumen de dichas actividades se han considerado los volúmenes anuales añadidos a la colección, tanto de monografías como de revistas obtenidas por compra, y que ha sido definida con el nombre de "COMPRAS".
 - Actividades relacionadas con el mantenimiento de las instalaciones y que tienen como objetivo el de ofrecer a los usuarios unas condiciones adecuadas de las mismas. Como variable hemos seleccionado el número de puntos de servicio que están disponibles para los usuarios y que se ha definido como "NBIBLIO".
- b) El de servicio a los usuarios, cuyo objetivo es el de dar respuesta a las necesidades de información que presentan los mismos en sus actividades de aprendizaje, docencia e investigación. Este macroproceso incluye los siguientes, procesos: información bibliográfica, consulta y recuperación de información; préstamo tanto en sala como a domicilio, e interbibliotecario; formación de usuarios; apoyo a la docencia; apoyo a la investigación y el de visibilidad de la producción científica de la universidad.

Como variables representativas de este apartado se han seleccionado:

- El número de préstamos domiciliarios, que se ha definido como "PRESTDOM".
- El número de préstamos interbibliotecarios que gestiona la biblioteca, tanto como centro solicitante como proveedor, y que se ha definido como "INTERB".

La variable "PERSONAL", que viene reflejada en este artículo, se ha construido como la suma del personal de plantilla de la biblioteca y los becarios, todos ellos reducidos a tiempo completo. En la tabla I se presentan los estadísticos descriptivos más relevantes pertenecientes a las 41 bibliotecas analizadas, pudiéndose observar que en el año 2012 se produce una reducción en el número de efectivos del personal, volumen de compras y préstamos interbibliotecarios, todo ello posiblemente debido a los efectos de la crisis económica.

4. RESULTADOS

4.1. Resumen de los resultados del cálculo de la eficiencia

En la Tabla II se muestra el resumen de los principales resultados de eficiencia, tanto bajo el supuesto de rendimientos constantes (CRS), como de rendimientos dependientes de la escala de funcionamiento (VRS) orientados a la minimización en el uso de inputs. Así mismo, se ha representado la eficiencia de escala de los años 2008 y 2012.

Como resultado más sobresaliente se puede señalar que el número de bibliotecas eficientes bajo el modelo CRS en 2008 es de 10 (24%), y bajo el modelo VRS es de 19 (46%). En el año 2012 estos valores corresponden a 8 bibliotecas eficientes en el modelo CRS (19%) y a 15 (36%) en el modelo VRS. El promedio de eficiencia de las 41 bibliotecas bajo el modelo CRS es de 0,79 para el

Tabla I. Resumen estadístico de las variables utilizadas en el estudio

Variables	Año 2008				Año 2012			
	Mínimo	Máximo	Media	D. T.*	Mínimo	Máximo	Media	D. T.*
PERSONAL	29	268	109,4	56,5	31	273	106,5	54,5
COMPRAS	3.268	43.425	17.845,1	9.780,2	2.247	52.360	11.194,6	9.132,4
ITEMS	127.653	1.446.411	587.204,9	296.412,2	171.901	1.895.421	710.646,6	370.869,8
NBIBLIO	1	21	10,5	6,4	1	22	9,9	5,9
PRESTDOM	58.082	637.061	245.477,2	163.171,7	51.343	767.285	259.235,1	177.938,9
INTERB	1.502	16.556	6.322,6	3.858,0	885	1.6256	4.657,0	3.302,3

D.T*. (Desviación típica)

Tabla II. Valores de eficiencia de las bibliotecas universitarias españolas en los años 2008 y 2012

Biblioteca	Año 2008			Año 2012		
	Eficiencia CRS	Eficiencia VRS	Eficiencia de Escala	Eficiencia CRS	Eficiencia VRS	Eficiencia de Escala
UDC	1,00	1,00	1,00	1,00	1,00	1,00
UA	0,51	0,53	0,95	0,47	0,47	1,00
UAH	1,00	1,00	1,00	1,00	1,00	1,00
UAL	0,44	0,54	0,81	0,48	0,64	0,75
UAB	0,71	1,00	0,71	0,64	0,78	0,81
UAM	0,79	1,00	0,79	0,92	1,00	0,92
UCA	0,98	1,00	0,98	0,91	0,91	0,99
UC	0,75	0,75	1,00	0,83	0,84	1,00
UC3M	1,00	1,00	1,00	0,77	0,88	0,87
UCLM	1,00	1,00	1,00	0,85	0,99	0,86
UCO	0,63	0,63	0,99	0,82	0,84	0,98
UEX	0,93	1,00	0,93	0,87	1,00	0,87
UdG	0,76	0,76	1,00	0,60	0,68	0,89
UGR	0,80	1,00	0,80	1,00	1,00	1,00
UHU	0,63	0,72	0,88	0,55	0,68	0,81
UIB	1,00	1,00	1,00	1,00	1,00	1,00
UJA	0,83	0,85	0,96	0,90	1,00	0,90
UJI	0,95	0,95	1,00	0,96	1,00	0,96
ULL	0,56	0,56	1,00	0,57	0,57	1,00
UR	1,00	1,00	1,00	1,00	1,00	1,00
ULPGC	0,74	0,75	0,99	0,84	0,85	0,99
ULE	0,91	0,91	1,00	1,00	1,00	1,00
UdL	0,81	0,81	1,00	0,63	0,78	0,81
UMA	0,59	0,59	1,00	0,57	0,58	0,99
UMH	1,00	1,00	1,00	1,00	1,00	1,00
UM	0,80	0,86	0,93	0,74	0,74	0,99
UO	0,70	0,80	0,88	0,74	0,77	0,96
UPO	1,00	1,00	1,00	1,00	1,00	1,00
UPV/EHU	1,00	1,00	0,99	0,83	1,00	0,83
UPC	0,44	0,45	0,98	0,50	0,51	0,98
UPM	0,47	0,59	0,79	0,46	0,57	0,81
UPV	1,00	1,00	1,00	0,82	0,84	0,98
UPF	0,92	1,00	0,92	0,82	1,00	0,82
UPN	0,67	0,68	0,98	0,51	0,56	0,90
URJC	0,62	0,62	0,99	0,69	0,75	0,92
URV	0,96	1,00	0,96	0,56	0,56	0,99
USC	0,60	0,99	0,61	0,61	0,90	0,67
US	0,62	1,00	0,62	0,66	1,00	0,66
UBU	0,70	0,81	0,86	0,79	0,91	0,86
UVA	0,73	0,78	0,94	0,78	0,78	0,99
UVI	0,67	0,69	0,96	0,60	0,64	0,94
Media	0,79	0,85	0,93	0,76	0,83	0,92

Nota: ver nombres de las universidades en Anexo.

año 2008 y 0,76 para el año 2012. Esto significa que, a nivel global, se podría reducir el personal en más de un 20%. En el caso de utilizar como elemento de evaluación el modelo VRS, la eficiencia media es 0,85 en el año 2008 y de 0,83 en el año 2012, lo que significa que, como mínimo, se podría reducir en un 15% el volumen del personal para conseguir su eficiencia. Por otro lado, los altos valores obtenidos de la eficiencia de escala indican que alrededor de la mitad de ese 15% de baja eficiencia se explica porque las bibliotecas en su conjunto no están trabajando con un tamaño óptimo. De hecho, de las 32 bibliotecas que presentan en el año 2008 un valor de eficiencia de escala inferior a 1, el 62,5%, es decir, 20 de ellas están operando con un retorno de escala decreciente; para el año 2012 este porcentaje baja al 44%, posiblemente debido a la reducción del personal que se produjo en las bibliotecas en ese periodo como resultado de la crisis económica y que se puede observar en la tabla I.

Hay que resaltar el comportamiento similar que tienen las bibliotecas en los años 2008 y 2012. Así, las 6 bibliotecas más ineficientes lo son en ambos años, alcanzando valores muy similares. Así mismo, todas las bibliotecas que son eficientes en 2012 también lo son en 2008, exceptuándose la biblioteca UGR que en el año 2012 incrementó el número de ítems informatizados. Sin embargo, existen 3 bibliotecas que eran eficientes en 2008, pero que en 2012 ya no lo fueron y cuya causa se debe, en parte, a una reducción en la compra de materiales bibliográficos. No obstante, aunque existan algunas diferencias de comportamientos individuales, los datos de eficiencia calculados en ambos años son en su conjunto muy similares, como se prueba por la existencia de una corre-

lación entre dichos valores del 0,789 para la eficiencia global y de 0,811 para la eficiencia técnica pura (significativos para un nivel de probabilidad del 99%). También es preciso señalar que la realización de la prueba no paramétrica de los rangos con signo de Wilcoxon para muestras relacionadas, nos indica que no se puede descartar la hipótesis de igualdad de los valores medios obtenidos de eficiencia para ambos años con un nivel de probabilidad del 95%.

4.2. Ranking de las bibliotecas

Para establecer un ranking entre las bibliotecas eficientes se han aplicado los tres métodos señalados en la sección de metodología (tablas III y IV). En el caso de la evaluación de la eficiencia cruzada, hemos utilizado el valor promedio de eficiencia que se obtiene cuando la biblioteca evaluada utiliza los pesos de las demás bibliotecas (evaluación por los pares), así como el índice de "maverick" para verificar su fiabilidad.

Como se puede observar en la tabla III, las bibliotecas UIB, y UDC, son las que mejor se comportan en los tres métodos empleados durante el año 2008. Además, el bajo valor del índice de "maverick" para ambas nos indica la fiabilidad de dichos valores de eficiencia global. Por estos motivos, ambas ocuparían el primer puesto. A continuación se situarían las bibliotecas UMH, y UPV/EHU, con valores de eficiencia cruzada de alrededor del 80% y con unos índices de "maverick" inferiores a 25. A partir de aquí habría que destacar a la UCLM y ya, en menor medida, a la UAH, y UPV. Es necesario comentar que la biblioteca UC3M tiene un elevado índice de "maverick" (47,71), lo que nos sugiere que su valor de eficiencia puede ser discutible.

Tabla III. Comparación entre el número de unidades de referencia (U.R.) de las bibliotecas eficientes y los coeficientes de super-eficiencia y eficiencia cruzada en el modelo CCR para el 2008

Biblioteca	U.R.	Super-eficiencia	Eficiencia cruzada	Índice de "maverick"
UDC	17	1,24	0,96	3,84
UAH	9	1,11	0,76	31,66
UC3M	1	1,05	0,68	47,71
UCLM	10	1,06	0,76	31,74
UIB	25	1,40	0,93	8,15
UR	16	1,32	0,72	39,58
UMH	6	1,12	0,81	24,27
UPO	4	1,02	0,73	37,11
UPV/EHU	4	1,01	0,82	21,27
UPV	2	1,11	0,76	32,41

Los resultados para el año 2012 se presentan en la tabla IV, pudiéndose observar que las bibliotecas que ocupan los dos primeros lugares en eficiencia siguen siendo la UDC y la UIB. A continuación se mantiene la UR y la UMH. Hay que señalar que en este año desaparece de entre los primeros puestos la biblioteca UPV/EHU, debido a que incrementó el personal, al mismo tiempo que se redujeron las compras y el préstamo interbibliotecario, así como el número de puntos de servicio. En este mismo periodo, la biblioteca UGR entra entre las mejores, y esto se explica por el incremento en más de un 50% de los ejemplares introducidos en la base de datos bibliográfica. Si tenemos en cuenta el valor del índice de "maverick", podemos observar que dicho índice para la UPO, alcanza el valor de 53,96, lo que sugiere que posiblemente su valor de eficiencia sea poco fiable.

Teniendo en cuenta los resultados de los años 2008 y 2012, podríamos establecer un ranking en el que las bibliotecas UDC y UIB, alcanzarían el primer puesto, seguidas a una cierta distancia por las bibliotecas UR y UMH y detrás de estas se situarían las bibliotecas UAH, ULE y UPO, aun cuando los valores de eficiencia de este último grupo sería discutible por los altos valores de sus índices de "maverick".

4.3. Organización de los recursos humanos

Teniendo en cuenta los datos del posible ranking de bibliotecas, lo primero que detallaremos son las características organizativas que poseen las bibliotecas situadas en el primer puesto del mismo. Así, desde la perspectiva de la organización de los recursos humanos, la UDC se caracteriza por disponer de una estructura central en la que se localiza la dirección y 4 unidades administrativas dedicadas a la gestión de los procesos comunes bibliotecarios: automatización, técnico, gestión de colecciones, referencia y acceso a los documentos.

Además, cuenta con 16 bibliotecas de centro con una dotación global de 75 personas (media de 4,7 personas por biblioteca).

La biblioteca UIB también dispone, como la anterior, de una estructura central amplia con 7 unidades técnicas gestionadas por 25 personas y 10 puntos de servicio con una dotación de 5,8 personas de media.

En contraste con las bibliotecas referidas más arriba, nos encontramos con dos categorías de bibliotecas que presentan una baja eficiencia (entre el 40 y 60%). El primer caso está constituido por bibliotecas como la UPC, la UPM, o la UMA, que se caracterizan por disponer de una estructura central de coordinación muy pequeña, dándole protagonismo a las bibliotecas de centro, lo que supone una estructura de gestión muy descentralizada en la que los procesos, como los técnicos o el préstamo interbibliotecario, suelen estar incluidos. Este hecho determina que las bibliotecas de centro dispongan de una dotación media superior a 9 personas sin contar a los becarios. El segundo tipo está constituido por bibliotecas como la UA, la cual dispone de una estructura central potente, con más de 50 personas, y una dotación media superior a 9 personas por biblioteca de centro.

Entre las bibliotecas que disponen de menos de 6 puntos de servicio debemos señalar, como ejemplo de eficiencia, la UMH que está organizada de forma descentralizada en tres secciones, las cuales dan servicio a las cinco bibliotecas existentes en los diferentes campus en los que se organiza la universidad. La dotación de personal de cada una de las secciones era, en 2008, entre 7 y 10 personas; al contrario, dentro de este grupo la que resulta más ineficiente es la biblioteca UAL, que dispone de una dotación de personal para un solo punto de servicio superior a la biblioteca UR con un solo punto de servicio, o la UMH con 5 puntos de servicio.

Tabla IV. Comparación entre el número de unidades de referencia (U.R.) de las bibliotecas eficientes, y los coeficientes de super-eficiencia y eficiencia cruzada en el modelo CCR para el 2012

Biblioteca	U.R.	Super-eficiencia	Eficiencia cruzada	Índice de maverick
UDC	26	1,40	0,95	5,44
UAH	5	1,07	0,71	40,85
UGR	13	1,23	0,75	33,88
UIB	27	1,43	0,91	9,75
UR	17	1,12	0,76	31,73
ULE	10	1,12	0,71	40,13
UMH	3	1,00	0,72	39,02
UPO	7	1,20	0,65	53,96

5. CONCLUSIONES

En este artículo hemos analizado las posibilidades de establecimiento de un ranking entre las bibliotecas universitarias españolas, que podría ser utilizado para identificar a las mejores bibliotecas en relación con la eficiencia de los procesos que lleva a cabo su personal. Como resultado relevante, debemos señalar que las bibliotecas, en su conjunto, presentan una eficiencia global y técnica alta en los dos años estudiados. No obstante, existen algunas bibliotecas con valores de eficiencia inferiores al 60% que deberían mejorar la organización de sus recursos humanos. Para llevar a cabo esta mejora, hemos considerado que sería importante la identificación, por un lado, de aquellas bibliotecas que podrían ser modelo de referencia y, por otro, conocer qué características serían claves en la organización de los recursos humanos para elevar su nivel de eficiencia.

Como primera conclusión, podríamos indicar que la evaluación de la eficiencia cruzada puede ser un método muy efectivo para la identificación de las bibliotecas más eficientes, dando unos resultados que son coherentes con los otros dos métodos utilizados en este trabajo pero que, además, permite a través del cálculo del índice de "maverick" conocer la fiabilidad del valor de eficiencia.

Como segunda conclusión destacamos la estabilidad de la eficiencia en el conjunto de las bibliotecas universitarias en los dos años analizados. La

utilidad del cálculo de la eficiencia global y de la eficiencia pura y de escala, nos permite conocer si la causa de la ineficiencia reside en un funcionamiento por debajo o por encima de su tamaño óptimo.

En tercer lugar, se resalta el interés que tiene el estudiar la relación existente entre los datos de eficiencia con la organización de las bibliotecas. En este caso, se ha comprobado que las bibliotecas más eficientes en el uso de sus recursos humanos se caracterizan, por un lado, en que disponen de una estructura centralizada en la que se realizan todos aquellos procesos que son comunes a todas las bibliotecas de centro, lo que permite aprovechar las ventajas de la automatización de dichos procesos. Por otro lado, estas bibliotecas disponen de un número de recursos humanos asignados a las bibliotecas de centro inferior a los que existen en las bibliotecas evaluadas como muy ineficientes.

Como conclusión última, se sugiere que las bibliotecas menos eficientes deberían, por un lado, centralizar los procesos que son transversales como el préstamo interbibliotecario o las compras, y por otro, centrar sus esfuerzos en reducir el personal en aquellos procesos que puedan ser gestionados electrónicamente, como es el caso del mismo préstamo interbibliotecario, o la introducción de herramientas electrónicas en otros procesos como el préstamo, la respuesta a consultas de los usuarios, etc.

6. REFERENCIAS

- Adler, N; Friedman, L; Sinuany-Stern, Z. (2002). Review of ranking methods in the data envelopment analysis context. *European Journal of Operational Research*, vol. 140, 249-265. [http://dx.doi.org/10.1016/S0377-2217\(02\)00068-1](http://dx.doi.org/10.1016/S0377-2217(02)00068-1)
- Andersen, P; Petersen, N. C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science*, vol. 39, 1261-1264. <http://dx.doi.org/10.1287/mnsc.39.10.1261>
- Angulo-Meza, L; Lins, M.P.E. (2002). Review of methods for increasing discrimination in data envelopment analysis, *Annals of Operations Research*, vol. 116, 225-242. <http://dx.doi.org/10.1023/A:1021340616758>
- Banker, R. D; Charnes, A; Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, vol. 30(9), 1078-1092. <http://dx.doi.org/10.1287/mnsc.30.9.1078>
- Banker, R. D; Chang, H. (2006). The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research*, vol. 175, 1311-1320. <http://dx.doi.org/10.1016/j.ejor.2005.06.028>
- Bellini, T. (2012). Forward search outlier detection in data envelopment analysis. *European Journal of Operational Research*, vol. 216, 200-207. <http://dx.doi.org/10.1016/j.ejor.2011.07.023>
- Bergendahl, G. (1998). DEA and benchmarks -an application to Nordic banks. *Annals of Operations Research*, vol. 82, 233-249. <http://dx.doi.org/10.1023/A:1018910719517>
- Carvalho, F.A; Jorge, M. J; Jorge, M. F; Russo, M; Sá, N. O. (2012). Library performance management in Rio De Janeiro, Brazil: applying DEA to a sample of university libraries in 2006-2007. *Library Management*, Vol. 33(4/5), 297-306. <http://dx.doi.org/10.1108/01435121211242335>
- Carvalho, F.A; Jorge, M. J; Jorge, M. F; Medeiros, R.O. (2013). Análise de eficiência e desempenho no longo prazo: ilustração empírica de um modelo para avaliação de bibliotecas públicas. *Investigación Bibliotecológica*, vol. 27 (60), 71-95.
- Charnes, A; Cooper, W. W; Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, vol. 2(6), 429-444. [http://dx.doi.org/10.1016/0377-2217\(78\)90138-8](http://dx.doi.org/10.1016/0377-2217(78)90138-8)

- Chen, T. (1997). A measurement of the resource utilization efficiency of university libraries. *International Journal of Production Economics*, vol. 53(1), 71-80. [http://dx.doi.org/10.1016/S0925-5273\(97\)00102-3](http://dx.doi.org/10.1016/S0925-5273(97)00102-3)
- Doyle, J. R; Green, R.H. (1994). Efficiency and cross-efficiency in DEA: Derivations, meanings and uses. *Journal of the Operational Research Society*, vol. 45, 567-578. <http://dx.doi.org/10.1057/jors.1994.84> / <http://dx.doi.org/10.2307/2584392>
- Hwang, M. J; Shieh, J. C; Hsieh, C. C. (2012). Using data envelopment analysis to evaluate library electronic databases. *Journal of Educational Media and Library Science*, vol. 49 (3), 343-368.
- Jarrillo Calvarro, S. (2012). Bibliotecas universitarias y costes. ¿Valemos lo que costamos? *Boletín de la Asociación Andaluza de Bibliotecarios*. vol. 103, 45-57.
- Johnson, A. L; McGinnis, L. F. (2008). Outlier detection in two-stage semiparametric DEA models. *European Journal of Operational Research*, vol. 187, 629-635. <http://dx.doi.org/10.1016/j.ejor.2007.03.041>
- Kao, C; Lin, Y-C. (1999). Comparing University Libraries of Different University Size. *Libri*, vol. 49, 150-158. <http://dx.doi.org/10.1515/libr.1999.49.3.150>
- Kao, C; Liu, S-T. (2000). Data envelopment analysis with missing data: an application to University libraries in Taiwan. *Journal of the Operational Research Society*, vol. 51, 897-905. <http://dx.doi.org/10.2307/254045> / <http://dx.doi.org/10.1057/palgrave.jors.2600056>
- Kao, C; Lin Y-C. (2004). Evaluation of the university libraries in Taiwan: total measure versus ratio measure. *Journal Operational Research Society*, vol. 55 (12), 1256-1265. <http://dx.doi.org/10.1057/palgrave.jors.2601786>
- Lai, M. C; Wang, W. K; Huang, H. C; Meng- Kao, M. C. (2011). Linking the benchmarking tool to a knowledge-based system for performance improvement. *Expert Systems with Applications*, vol. 38, 10579-10586. <http://dx.doi.org/10.1016/j.eswa.2011.02.101>
- Lee, C. W; Kwak, N.K.; Garrett, W. A. (2013). A Comparative Appraisal of Operational Efficiency in U.S. Research-University Libraries: A DEA Approach. En: Lawrence, K. D; Kleinman, G (editores) *Applications of Management Science*, vol. 16, pp. 117-130, Emerald Group Publishing Limited.
- Liu, S. T; Chuang, M. (2009). Fuzzy efficiency measures in fuzzy DEA/AR with application to university libraries. *Expert Systems with Applications*, vol. 36, 1105-1113. <http://dx.doi.org/10.1016/j.eswa.2007.10.013>
- Lu, W. M; Lo, S. F. (2009). An interactive benchmark model ranking performers- Application to financial holding companies. *Mathematical and Computer Modelling*, vol. 49, 172-179. <http://dx.doi.org/10.1016/j.mcm.2008.06.008>
- Mann, G. M. (1997). *Efficiency evaluations of North American university libraries: Data envelopment analysis*. Doctoral Thesis. Montreal, Canada: McGill University. <http://www.collectionscanada.gc.ca/obj/s4/f2/dsk2/ftp01/MQ29750.pdf>. [consulta: 4 de junio de 2014].
- Noh, Y. (2011). Evaluation of the resource utilization efficiency of university libraries using DEA techniques and a proposal of alternative evaluation variables. *Library Hi Technology*, vol. 29 (4), 697-724. <http://dx.doi.org/10.1108/07378831111189787>
- Owadie, S. (2011). *Measuring the efficiency of university libraries, Ghana-Using data envelopment analysis*. Tesis Doctoral. University of Science and Technology-Ghana.
- REBIUN. Red de Bibliotecas Universitarias. Estadística REBIUN. <https://rebiun.um.es/rebiun/admin/ManageIndicatorsPage;jsessionid=0BF061E983FD83761AF6AB7856849971> [consultada desde el 10 de julio hasta el 20 diciembre de 2014].
- Reichmann, G. (2004). Measuring university library efficiency using data envelopment analysis. *Libri*, vol. 54 (2), 136-46. <http://dx.doi.org/10.1515/libr.2004.136>
- Reichmann, G; Sommersguter-Reichmann, M. (2006). University library benchmarking: An international comparison using DEA. *International Journal of Production Economics*, vol. 100, (1), 131-147. <http://dx.doi.org/10.1016/j.ijpe.2004.10.007>
- Reichmann, G; Sommersguter-Reichmann, M. (2010). Efficiency measures and productivity indexes in the context of university library benchmarking. *Applied Economics*, vol. 42 (3), 311-23. <http://dx.doi.org/10.1080/00036840701604511>
- Saunders, E. S. (2003). Cost efficiency in ARL academic libraries. *The Bottom Line*, vol. 16(1), 5-14. <http://dx.doi.org/10.1108/08880450310464009>
- Sexton, T.R; Silkman, R. H; Hogan, A. J. (1986). Data envelopment analysis: critique and extensions. In: Silkman, R.H. (ed.), *Measuring efficiency: an assessment of data envelopment analysis*. *New Directions for Program Evaluation*, vol. 1986 (32), 73-105. <http://dx.doi.org/10.1002/ev.1441>
- Shahwan, T. M; Kaba, A. (2013). Efficiency analysis of GCC academic libraries: an application of data envelopment analysis. *Performance Measurement and Metrics*, vol. 14(3), 197-210. <http://dx.doi.org/10.1108/PMM-07-2013-0023>
- Shim, W. (2000). Assessing technical efficiency of research libraries. *Advances in Library Administration and Organization*, vol. 17, 243-339. [http://dx.doi.org/10.1016/s0732-0671\(00\)80011-5](http://dx.doi.org/10.1016/s0732-0671(00)80011-5)
- Shim W. (2003). Applying DEA technique to library evaluation in academic research libraries. *Library Trends*, vol 51(3), 312-32.

- Simar, L. (2003). Detecting Outliers in Frontiers Models: A Simple Approach. *Journal of Productivity Analysis*, vol. 20, 391-424. <http://dx.doi.org/10.1023/A:1027308001925>
- Simon Blas, C; Arias Coello, A; Simon Martin J. (2007). Aplicación de la técnica DEA en la medición de la eficiencia de las bibliotecas de la Universidad Complutense de Madrid. *Revista Española de Documentación Científica*, vol. 30 (1), 9-23.
- Simon, J; Simon, C; Arias, A. (2011). Changes in productivity of Spanish university libraries. *Omega*, vol. 39 (5), 578-588. <http://dx.doi.org/10.1016/j.omega.2010.12.003>
- Stancheva, N; Angelova, W. (2008). Measuring the efficiency of university libraries using data envelopment analysis. En *Research methodology in data envelopment analysis*. Ed. J. Kumar Mantri. pgs. 13-21. Universal Publisher. USA 2008.
- Talluri, F; Sarkis, J. (1997). Extensions in efficiency measurement of alternate machine component grouping solutions via Data Envelopment Analysis. *IEE Transactions on Engineering Management*, vol. 44(3), 299-304. <http://dx.doi.org/10.1109/17.618171>
- Torgersen, A. M; Førsund, F. R; Kittelsen, S.A.C. (1996). Slack-adjusted efficiency measures and ranking of efficient units. *Journal of Productivity Analysis*, vol. 7, 379-398. <http://dx.doi.org/10.1007/BF00162048>

ANEXO

Acrónimo	Nombre de las universidades	Acrónimo	Nombre de las universidades
UDC	Universidad de A Coruña	ULE	Universidad de León
UA	Universidad de Alicante	UdL	Universitat de Lleida
UAH	Universidad de Alcalá de Henares	UMA	Universidad de Málaga
UAL	Universidad de Almería	UMH	Universidad Miguel Hernández
UAB	Universitat Autònoma de Barcelona	UM	Universidad de Murcia
UAM	Universidad Autónoma de Madrid	UO	Universidad de Oviedo
UCA	Universidad de Cádiz	UPO	Universidad Pablo Olavide
UC	Universidad de Cantabria	UPV/EHU	Universidad del País Vasco
UC3M	Universidad Carlos III de Madrid	UPC	Universitat Politècnica de Catalunya
UCLM	Universidad de Castilla-La Mancha	UPM	Universidad Politécnica de Madrid
UCO	Universidad de Córdoba	UPV	Universitat Politècnica de Valencia
UEX	Universidad de Extremadura	UPF	Universitat Pompeu Fabra
UdG	Universitat de Girona	UPN	Universidad Pública de Navarra
UGR	Universidad de Granada	URJC	Universidad Rey Juan Carlos
UHU	Universidad de Huelva	URV	Universidad Rovira i Virgili
UIB	Universitat de les Illes Balears	USC	Universidade de Santiago de Compostela
UJA	Universidad de Jaén	US	Universidad de Sevilla
UJI	Universitat Jaume I	UBU	Universidad de Burgos
ULL	Universidad de La Laguna	UVA	Universidad de Valladolid
UR	Universidad de la Rioja	UVI	Universidad de Vigo
ULPGC	Universidad de Las Palmas de Gran Canaria		



ESTUDIOS / RESEARCH STUDIES

Evaluación de la accesibilidad web de los portales del Estado en Perú

Carlos Antonio Sam-Anlas* y Yudayly Stable-Rodríguez**

*Biblioteca Nacional del Perú.

**Instituto de Información Científica y Tecnológica. La Habana. Cuba.
Correo-e: csam@bnp.gob.pe, yuly@idict.cu

Recibido: 08-07-2014; 2ª versión: 15-05-2015; Aceptado: 01-06-2015.

Cómo citar este artículo/Citation: Sam_Anlas, C. A. y Stable-Rodríguez, Y. (2016). Evaluación de la accesibilidad web de los portales del Estado en Perú. *Revista Española de Documentación Científica*, 39(1): e120. doi: <http://dx.doi.org/10.3989/redc.2016.1.1213>

Resumen: En este artículo se presentan los resultados de un estudio para evaluar la accesibilidad de la página principal de algunos portales del Estado en Perú. La evaluación de la accesibilidad se ha realizado a partir del cumplimiento de las prioridades de la accesibilidad web contempladas en la legislación peruana, así como en las iniciativas de accesibilidad (WAI) establecidas por el World Wide Web Consortium (W3C), en la recomendación WCAG 1.0 y WCAG 2.0. Para tal fin, se identificaron cinco portales los cuales contaban con mayor acceso e importancia en el momento de la investigación y con la utilización de herramientas manuales y automáticas, se obtuvieron los resultados, que muestran las barreras y deficiencias en el acceso a la información a personas que tienen algún tipo de discapacidad.

Palabras clave: Accesibilidad web; discapacidad; WCAG 2.0; herramientas de evaluación.

Evaluating accessibility in Peruvian Government websites

Abstract: This paper presents the results of a study evaluating the accessibility of some major portals of the Peruvian Government. The accessibility evaluation was made in compliance with web accessibility priorities addressed in Peruvian law, as well as the guidelines established by the World Wide Web Consortium (W3C) in its WCAG 1.0 and WCAG 2.0 recommendations. To this aim, five portals of greatest importance and with highest levels of access during the period of investigation were identified. Manual and automatic tools were used to obtain the results. These revealed the barriers and gaps in information access that were encountered by people with disabilities.

Keywords: Web accessibility; disabilities; WCAG 2.0; assessment tools.

Copyright: © 2016 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0.

1. INTRODUCCIÓN

La accesibilidad web es una métrica de la Ingeniería Web que trata de elementos relacionados con la codificación y la presentación de información en el diseño y funcionalidad de un sitio web, que va a permitir que las personas con algún tipo de discapacidad puedan percibir, entender, navegar e interactuar de forma efectiva en la web, así como crear y aportar contenido. Su incumplimiento ocasiona discriminación debido a que no pueden utilizar la información disponible en la web de manera normal.

En la actualidad, no existe una definición formal y totalmente aceptada del concepto de accesibilidad web, sin embargo, la necesidad de que la Web sea universal y accesible para cualquier persona está presente desde sus inicios, debido a que era un requisito contemplado en su diseño por su creador Tim Berners-Lee.

Coincidiendo con Zubillaga del Río (2010), las definiciones sobre accesibilidad desde una perspectiva tecnológica son numerosas: “[...] se puede definir la accesibilidad Web como la posibilidad de que un producto o servicio web pueda ser accedido y usado por el mayor número posible de personas, indiferentemente de las limitaciones propias del individuo o de las derivadas del contexto de uso” (Hassan Montero y Martínez Fernández, 2003).

“Accesibilidad significa proporcionar flexibilidad para acomodarse a las necesidades de cada usuario y sus preferencias” (Segovia, 2007).

“Hablar de Accesibilidad Web es hablar del acceso de todos a la Web, independientemente del tipo de hardware, software, infraestructura de red, idioma, cultura, localización geográfica y capacidades de los usuarios” (W3C, 2011).

En resumen, principios como acceder, explorar, usar, flexibilidad, necesidades del usuario, preferencias, convergen en todas las definiciones anteriores, con independencia del modo de acceso y uso de la red y sus posibilidades.

La accesibilidad web no solo beneficia a las personas que presentan algún tipo de discapacidad, sino también a otros grupos de usuarios que presentan dificultades relacionadas con el envejecimiento o derivadas de situaciones desfavorables.

Por lo anterior, es importante que los sitios web cumplan con estándares de accesibilidad para lograr que un mayor número de usuarios los visiten, estándares que en muchos países han sido desarrollados con la anuencia del Estado, por ello, el objetivo de esta investigación consistió en evaluar la accesibilidad web de los principales portales del Estado en Perú.

2. ANTECEDENTES MÁS RECIENTES EN EL ESTUDIO DE LA ACCESIBILIDAD WEB

Aunque existe una gran cantidad de trabajos de investigación sobre accesibilidad web, se han seleccionado las investigaciones realizadas sobre accesibilidad web en portales de instituciones públicas (Tabla I), para conocer el comportamiento del tema, en instituciones similares, conociendo técnicas y resultados de las mismas, que pueden servir de base para el caso particular de Perú.

Estas investigaciones se caracterizan por la aplicación de diferentes metodologías para la evaluación de la accesibilidad, pero en algunos casos aún no han sido suficientemente sistematizadas. Por otra parte, en América Latina, los reportes de compilaciones que reúnen las legislaciones de los países para viabilizar el tema de accesibilidad web (PNUD, 2009; Rosales, 2011), poseen información dispersa y con falta de sistematicidad.

3. ESTÁNDARES DE ACCESIBILIDAD WEB

Existen varias normas y estándares en el diseño de páginas web encaminadas a facilitar la consecución de un adecuado nivel de accesibilidad (Serrano Mascaraque y otros, 2010).

Estas normas y estándares han sido desarrolladas por diferentes organizaciones. Cuando se trata de organizaciones reconocidas por los gobiernos como de estandarización, estas normas son estándares oficiales, como es el caso de los estándares de la Organización Internacional de Normalización (ISO), la Asociación Española de Normalización y Certificación (AENOR), el Instituto de Estándares Nacional Americano (*American National Standards Institute-ANSI*), y el Instituto Británico de Estándares (*British Standards Institute - BSI*).

Cuando son recomendaciones de otras organizaciones que se han convertido en verdaderos estándares por su uso en la práctica, se denominan estándares “de facto”, como es el caso de los estándares de la *World Wide Web Consortium* (W3C), el Instituto Europeo de Estándares de Telecomunicación (*European Standards Telecommunications Institute - ETSI*), la Sociedad de Factores Humanos y Ergonomía (*Human Factor and Ergonomics Society - HFES*), así como el Consorcio IMS Global de Aprendizaje (IMS Global Learning) (Varela y otros, 2012).

En la actualidad, en el ámbito internacional las recomendaciones del W3C-WAI constituyen la referencia en cuanto a criterios y estrategias de accesibilidad a Internet, con el objetivo de establecer recomendaciones para conseguir contenidos, na-

Tabla I. Trabajos sobre evaluación de accesibilidad web en instituciones públicas

Investigación	Autores
Evaluación de la Accesibilidad y usabilidad del Sitio Web de la Unión Cuba-Petróleo	(Montano, 2014)
Dificultades en la Accesibilidad Web de las Universidades Españolas de acuerdo a la Norma WCAG 2.0	(Chacón-Medina y otros, 2013)
Evaluación de la accesibilidad de páginas web de universidades españolas y extranjeras incluidas en rankings universitarios internacionales	(Hilera y otros, 2013)
Evaluación de la accesibilidad de portales web en instituciones educativas en el área de Centroamérica	(Hernández y Amado-Salvatierra, 2012)
Evaluación de la sede web del Centro Regional de Información sobre Desastres (CRID): aplicación del cuestionario BRACAD	(Medina, 2012)
Accessibility and readability of university websites in Finland	(Karhu y otros, 2012)
Estudio de accesibilidad Web en las bibliotecas virtuales en salud en Perú	(Quiliano y Ríos, 2012)
Estudio comparativo de accesibilidad Web en portales informativos de universidades peruanas de educación a distancia	(Córdova, 2012)
Análisis de accesibilidad de los sitios web de noticias de las provincias del Nordeste Argentino	(Acevedo y otros, 2011)
La accesibilidad como elemento del proceso educativo: análisis del modelo de accesibilidad de la universidad Complutense de Madrid para atender las necesidades educativas de los estudiantes con discapacidad	(Zubillaga del Rio, 2010)
Evaluación de Accesibilidad de Sitios Web oficiales: Ministerios y Municipios de la Provincia de Buenos Aires	(Díaz y otros, 2008)
Evaluación comparativa de la accesibilidad de los espacios Web de las bibliotecas universitarias españolas y norteamericanas	(Caballero- Cortés y otros, 2008)

vegadores, y entornos de desarrollo de web accesibles. Coincidiendo con Varela y otros (2012), los estándares de W3C-WAI (*Web Accessibility Initiative del World Wide Web Consortium*) son los más importantes en el ámbito de la accesibilidad web, y todos son de acceso gratuito en la Web, disponibles para descarga directa.

Entre sus recomendaciones se encuentran las *Web Content Accessibility Guidelines* (WCAG), o conjunto de pautas para conseguir páginas web accesibles. La última versión de esta recomendación es WCAG 2.0 (W3C, 2011), adoptada en 2012 por ISO como estándar (ISO, 2012), que establece un total de 12 pautas.

4. PAUTAS DE ACCESIBILIDAD WEB DE LA WCAG (WEB CONTENT ACCESSIBILITY GUIDELINES)

Las pautas WCAG están dirigidas a quienes generan contenidos para la Web. Consisten en recomendaciones puntuales, redactadas de una forma menos genérica que el resto de las pautas anteriores. El foco está puesto en lograr que el contenido sea presentado de forma accesible. La iniciativa de Accesibilidad Web (WAI) del World Wide Web Consortium (W3C) elaboró guías anexas que explotan cada punto y detallan los pasos a seguir para implementarlas y cumplirlas (Naftali, 2010).

Las recomendaciones del W3C-WAI (*Web Accessibility Initiative del World Wide Web Consortium*) en el ámbito internacional, constituyen referencias en cuanto a criterios y estrategias de accesibilidad a Internet, no son normas estrictas, pero indican lo que el usuario debe poder hacer y qué tipo de información debe estar disponible. De esta forma se pueden realizar consultas y utilizar servicios conforme a las actuales tecnologías, y participar activamente en la sociedad de la información.

Al adoptar las pautas se permitirá a las personas con discapacidad, utilizar los mismos servicios en línea (sanidad en línea, administración en línea y educación en línea), como los demás ciudadanos. Estas pautas se han desarrollado en dos versiones, las cuales se describen a continuación.

La versión 1.0 de las Pautas de Accesibilidad al Contenido en la Web (*Web Content Accessibility Guidelines 1.0*) fue un avance importante para lograr que Internet sea más accesible para las personas con discapacidad. Finalizadas en 1999, las WCAG 1.0 proporcionan 14 directrices y varios puntos de verificación (65) que pueden utilizarse para determinar si se cumple la accesibilidad de una página web, mediante tres prioridades, niveles de cumplimiento o niveles de adecuación (la medida en que una página web cumple las directrices) (W3C, 1999), las cuales son:

- **Prioridad 1:** Debe ser cumplido. Es un requerimiento básico. Hace referencia a aquellos requisitos de los puntos de verificación que una página Web debe cumplir para no dificultar o impedir el acceso a un gran grupo de usuarios.
- **Prioridad 2:** Debería ser cumplido. Remueve ciertas barreras. O sea, se agrupan aquellos elementos que "deben" satisfacerse para eliminar importantes barreras en el acceso por parte de diferentes grupos.
- **Prioridad 3:** Podría ser cumplido. Mejoraría la accesibilidad para ciertos grupos. Hace referencia a ciertas características que debe cumplir la página Web para reducir las dificultades que pueden encontrar algunos grupos de usuarios en su acceso. Se trata de recomendaciones que afectan a un número menor de usuarios potenciales, pero que mejoran la accesibilidad y usabilidad de la página en general.

Del conjunto de prioridades en el sitio, las 14 pautas definen tres niveles de conformidad (W3C, 2011):

- **Nivel A:** Todos los puntos de control de prioridad 1 son cumplidos.
- **Nivel AA:** Todos los puntos de control de prioridades 1 y 2 son cumplidos.
- **Nivel AAA:** Todos los puntos de control de prioridades 1, 2 y 3 son cumplidos.

Las WCAG 1.0 se volvieron obsoletas, respecto al avance de las tecnologías web y las tecnologías para las personas con discapacidad, siendo compleja su adecuación, debido a que algunos puntos de control fueron menos importantes y más complejos de comprobar, lo que conllevó al desarrollo de la versión 2.0, conocida como WCAG 2.0.

Las Pautas de Accesibilidad al Contenido en la Web 2.0 (Web Content Accessibility Guidelines 2.0) se fundamentan en WCAG 1.0, pero también introducen algunos cambios significativos (W3C, 2011).

El cambio de pautas dio lugar a un número reducido de ideas de nivel superior o *principios*. WCAG 1.0 tenía catorce principios en el nivel superior, sin embargo, WCAG 2.0 sitúa únicamente cuatro principios en el nivel superior en virtud de los cuales se organizan pautas más específicas con sus correspondientes puntos de verificación (61), llamadas *criterios de éxito*. Cada uno de estos cuatro principios se indican como:

- **Perceptible:** se debe poder percibir.
- **Operable:** ha de ser posible manipular y manejar los elementos de interacción del contenido.

- **Comprensible:** tanto el contenido como los controles para su manejo han de ser comprensibles.
- **Robusto:** debe ser suficientemente robusto como para funcionar con tecnologías actuales y futuras.

Las pautas son de hecho una norma mundialmente reconocida para la creación de sitios web accesibles.

5. NORMATIVAS LEGALES SOBRE ACCESIBILIDAD WEB

En algunos países, algunas leyes no tienen el alcance de aplicación de las normas de accesibilidad web como lo hacen otras leyes en países como Francia donde se estableció que los servicios de comunicación al público en línea del Gobierno, de las autoridades locales y de las instituciones públicas deben ser accesibles a las personas con discapacidad.

La incorporación de la accesibilidad web en la legislación de los países de América Latina es bastante reciente y desigual (Díez Canseco, 2009; Huerta Peralta, 2006; Rosales, 2011), a pesar de que toman en cuenta condiciones básicas para el acceso de las personas con discapacidad a las tecnologías, productos y servicios relacionados con la sociedad de la información, sumado a que los gobiernos de estos países son signatarios y han ratificado la Convención sobre los Derechos de las Personas con Discapacidad (ONU, 2008).

Una de las primeras legislaciones sobre accesibilidad web que se promulgan en la región, fue en Argentina en noviembre de 2010 mediante la *Ley 26.653 de accesibilidad de la información en las páginas web*. En esta Ley se entiende por accesibilidad a "la posibilidad de que la información de la página Web, comprendida y consultada por personas con discapacidad y por usuarios que posean diversas configuraciones en su equipamiento o en sus programas". Posterior a ello, el 27 de junio de 2011, se publica la Resolución 69/2011, que aprueba la "Guía de Accesibilidad para Sitios Web del Sector Público Nacional". No obstante, aunque esta resolución es de 2011, la guía se basa en WCAG 1.0 en vez de WCAG 2.0, cuando esta última fue aprobada en diciembre de 2008.

La accesibilidad Web en el Perú, tiene sus orígenes en el ingreso de Internet en el país. A partir del 2000, se comenzó a regular algunos de estos temas con la Ley de Firmas y Certificados Digitales mediante la Ley N° 27269, la cual tiene por objeto regular la utilización de la firma electrónica otorgándole la misma validez y eficacia jurídica que el uso de una manuscrita u otra análoga, pero no

quedó bien definido como se va a introducir este proceso en la sociedad y especialmente para las personas con discapacidad visual.

En el 2005, se promulgó la Ley N° 28530 - Ley de Promoción de Acceso a Internet para Personas con Discapacidad y de adecuación del espacio físico en cabinas públicas de Internet, se destaca porque no sólo obliga a las entidades públicas y universidades a ofrecer sitios Web accesibles, sino también a las personas naturales o jurídicas que presten servicios de información al consumidor y otros servicios a través de páginas Web. Sin embargo, esta Ley tiene un problema grave debido a que sólo hace referencia a las personas con discapacidad visual, olvidando otros grupos de usuarios. Tampoco indica los niveles de accesibilidad que se deben cumplir.

El 25 de marzo de 2009 se decretó la Resolución Ministerial 126-2009-PCM de la Presidencia del Consejo de Ministros, que aprobó los lineamientos para accesibilidad a páginas Web y aplicaciones para telefonía móvil para instituciones públicas del Sistema Nacional de Informática. El propósito de este documento es que las instituciones públicas del Estado peruano apliquen las "Pautas de Accesibilidad de los Contenidos Web 1.0" (WCAG1.0) para mejorar sus páginas web y crear las bases para el desarrollo de aplicaciones para equipos móviles (telefonía celular), en el marco de la sociedad de la información, la inclusión social y con ello la accesibilidad a las personas con discapacidad y otros.

La Resolución Ministerial 126-2009-PCM de la Presidencia del Consejo de Ministros, es aún deficiente en materia de accesibilidad coincidiendo con Córdova Solís, 2012, debido a que no solo se debió establecer que se tienen que aplicar las pautas de accesibilidad del contenido en la web 1.0, sino también las pautas web 2.0, de forma paulatina. Por otra parte, la resolución está dirigida sólo para las páginas web de las instituciones públicas del Sistema Nacional de Informática.

6. MÉTODOS Y MATERIALES

Para la evaluación de la accesibilidad web se han utilizado las herramientas disponibles que permiten automatizar este proceso, según las pautas WCAG 1.0 y 2.0. No obstante, la evaluación de la accesibilidad requiere que la web en cuestión sea revisada mediante pruebas manuales por un usuario con discapacidad, mediante cuestionarios de forma presencial, o mediante entrevistas donde describan sus testimonios o experiencias concretas en cuanto a la accesibilidad.

En este estudio, el proceso de evaluación de la accesibilidad web seguido ha constado de dos fases:

Fase 1. Validación automática de la página inicial. Utilizando las herramientas específicas que permiten esbozar una primera idea del grado de accesibilidad de una página web, detectando los principales errores de la página principal como de otras páginas de la web.

A partir de los elementos considerados por Hilera y otros, 2013 en esta investigación también se han seleccionado otras páginas representativas de los 5 portales objeto de análisis:

- ✓ Una página que contenga formularios de visita habitual por los usuarios.
- ✓ Una página que contenga tablas.

Lo anterior tiene como objetivo contrastar que la accesibilidad de las páginas de inicio de los cinco portales no es diferente a las páginas interiores de esos, debido a que las páginas internas son utilizadas por los usuarios para realizar gestiones, o utilizar los formularios asociados o consultar tablas, y si estos no se han diseñado de forma adecuada, pueden presentar problemas de accesibilidad.

Se seleccionaron las métricas que utilizaron otros autores (Hernández y Salvatierra, 2012; Hilera y otros, 2013; Karhu y otros, 2012), las cuales se derivan de las pautas WCAG 2.0, sintetizadas en un conjunto de verificaciones técnicas de los aspectos más relevantes y con mayor incidencia. Las verificaciones se reúnen en un conjunto de 14 indicadores referidos a las recomendaciones WCAG los cuales son: Documentos web válidos, Imágenes, Encabezados, Enlaces, Contraste y uso semántico de los colores, Presentación, Tamaño del texto, Formularios, Tablas de datos, Accesibilidad a través de teclado, Destellos, Navegable, Comprensible, Tiempo suficiente.

Para evaluar cada uno de estos puntos se tuvo en cuenta uno o varios criterios de éxito de las WCAG 2.0, utilizados por Hilera y otros (2013), los cuales lo sustentaron en el estudio realizado por INTECO (2008). Evaluando también las comprobaciones en base a los valores "Bien", "Mal", "Pocos fallos" y "No aplicable (NA)", y calculando el porcentaje de éxito de la página con la siguiente expresión:

$$\text{Porcentaje de éxito} = 100\% \times \text{Bien} + 50\% \times \text{Pocos fallos} / \text{Total indicadores aplicables}$$

El total de indicadores aplicables se obtiene del número de indicadores evaluados, siendo 14, y el Número de páginas evaluadas, que es 2. Por lo tanto, el total de indicadores en la evaluación de los portales del Estado es de 28, eliminando pre-

viamente los indicadores no aplicables (NA), por lo que en cada una de las páginas este número tomará un valor distinto.

Fase 2. Validación manual. Con el fin de detallar y corregir la validación automática, se utilizan herramientas que simulan el entorno real de utilización accesible de los contenidos de la web, para revisar aquellos elementos determinados en la validación automática y comprobar cuales elementos no han sido tratados, por depender del contexto la semántica, entre otros, que deben ser validados manualmente siguiendo un conjunto de criterios.

En la investigación la evaluación se realizó en el primer semestre del 2014 a los portales más citados en el proyecto de gobierno electrónico del Perú (tabla II), utilizando las pautas y puntos de verificación de la WACG 1.0 y WACG 2.0, para los 3 niveles de adecuación.

Utilizando HERA se evaluaron las WCAG 1.0, y con el validador TAW, se evaluaron las WCAG 1.0 y las WCAG 2.0. El análisis se dividió en los tres niveles de prioridad que existen, atendiendo al grado de repercusión que el incumplimiento de sus puntos de verificación afecta a los diferentes usuarios.

Mediante la Barra de Herramientas de Accesibilidad Web, proporcionada por Accessible Information Solutions (AIS) de Vision Australia, traducida por Technosite (Grupo Fundosa (Fundación ONCE), Madrid, España), se realizó la evaluación manual, con el objetivo de revisar los elementos no tratados en la validación automática. Esta barra realiza funciones similares a la *Web Developer Toolbar Plugin* para *Mozilla Firefox*, como: redimensionar el tamaño del navegador, activar/desactivar las hojas de estilo, localizar características desaconsejadas, analizar la página con los validadores HTML o CSS, listar o resaltar las imágenes presentes o sustituirlas por su texto alternativo, hacer pruebas de color, identificar y resaltar los elementos estructurales o semánticos de la página Web (encabezados, listas, tablas de datos, *frames*).

La evaluación manual genera informes que incluyen una lista de los problemas encontrados, con referencia a su posición en el código de la página, eliminando falsos negativos y falsos positivos. Por ello, a partir de Zubillaga del Río (2010) se sintetizaron los criterios de análisis para proporcionar una visión sintética de las pautas de accesibilidad de un sitio Web, evaluado con herramientas manuales, a partir de los siguientes:

- 1. La utilización de HTML y CSS válidos:** Este aspecto establece que tanto el código HTML empleado en las páginas como el código de las hojas de estilo deben estar correctamente expresados y validados por las gramáticas formales; en este caso según las especificaciones HTML, XHTML y CSS2. Estos errores de código provocan que la visualización de la página sea diferente en función del navegador que se utilice, debido a que existen elementos no soportados por todos los navegadores. Según el W3C, un código HTML correcto asegura una compatibilidad total con cualquier navegador. La mayoría de los errores debidos a un uso incorrecto de HTML se deben a la mala utilización de atributos y elementos.
- 2. La navegación accesible (encabezados, enlaces y listas):** Los encabezados o títulos de sección, sirven para marcar la estructura de la información en cada una de las páginas. Se deben utilizar de forma correcta, respetando en todo momento el nivel de profundidad. Por ejemplo, no se debería utilizar, sin justificación, una etiqueta de encabezado de nivel 1 (h1) y seguidamente una de nivel 3 (h3). Por otra parte, los enlaces constituyen el componente más importante de un sitio web, debido a que permiten al usuario navegar entre páginas y seleccionar los contenidos a los que desea acceder, o sea, permite que el texto o la imagen que forman un enlace cumplan su objetivo de forma autónoma. Esto quiere decir que el enlace debe proporcionar al usuario in-

Tabla II. Portales seleccionados para el análisis de la accesibilidad web en Perú

Nombre del Portal	URL
Portal del Estado Peruano	http://www.peru.gob.pe
Portal del Consejo Nacional para la Integración de la Persona con Discapacidad (Conadis)	http://www.conadisperu.gob.pe/
Portal Servicios al Ciudadano	http://www.serviciosalciudadano.gob.pe
Portal del Ministerio de la Mujer y Poblaciones Vulnerables	http://www.mimp.gob.pe
Ministerio de Desarrollo e Inclusión Social	http://www.midis.gob.pe/

dicios claros sobre qué encontrará en la página que se mostrará luego de seleccionarlo.

- 3. La descripción de las imágenes.** Las imágenes complementan una información textual o simplemente hacen que las páginas sean más agradables, debido a que los usuarios, por varias causas (personas con discapacidad visual, los que utilizan navegadores de sólo texto, o aquellos que cancelan las descargas de imágenes por lentitud en la conectividad), no perciben las imágenes, siendo fundamental que estas ofrezcan una alternativa, para que transmitan información relevante para el uso de los contenidos web. Por ejemplo, los lectores de pantalla que utilizan las personas ciegas para acceder a Internet y los navegadores de sólo texto, muestran si existe el texto alternativo; en caso contrario, muestran el nombre del fichero de la imagen, lo cual puede resultar molesto para el usuario.
- 4. La estructura de las tablas de datos:** Las tablas deben señalar debidamente los encabezados de fila y columna o usar marcadores para asociar las celdas de encabezamiento con las celdas de datos en las tablas con dos o más niveles lógicos de encabezado. El marcado debe seguir las especificaciones de W3C. En caso contrario, un lector de pantalla no podrá ofrecer al usuario la información que necesita para relacionar los datos de la tabla. No se deben marcar los encabezados de fila o columna modificando el contenido de las mismas (por ejemplo, utilizando texto en negrita), sino marcando con el atributo «th» para que sean interpretados correctamente por los navegadores como tales encabezados. Además, para las tablas de datos que tienen dos o más niveles lógicos de encabezados de fila o columna se deben utilizar marcadores para asociar las celdas de encabezado y las de datos.

7. RESULTADOS Y DISCUSIÓN

7.1. Resultados de la evaluación de la fase 1

En la tabla III, se muestran los resultados para cada una de las tres prioridades, utilizando las herramientas Hera y TAW.

El nivel de prioridad 1 es de cumplimiento obligatorio para los portales del Estado del Perú según la Resolución Ministerial 126-2009-PCM de la Presidencia del Consejo de Ministros. Sin embargo, su revisión automática indicó dificultad, debido a que existen 43 criterios que deben ser verificables y solo el Portal del Estado peruano mostró un solo elemento evaluado como bien.

Algunas de las dificultades que como resultado arrojó el validador Hera son:

- **Punto de verificación 1.1:** Está relacionado con proporcionar un *texto equivalente para todo elemento no textual*. Esto incluye: imágenes, representaciones gráficas del texto (incluyendo símbolos), áreas de mapas de imagen, animaciones (por ejemplo, GIFs animados), "applets" y objetos programados, "ASCII art", marcos, scripts, imágenes usadas como viñetas en las listas, espaciadores, botones gráficos, sonidos (ejecutados con o sin la interacción del usuario), archivos exclusivamente auditivos, pista sonora del vídeo y vídeos.
- **Punto de verificación 4.1:** Relacionado con los cambios en el idioma, donde se debe verificar que todos los cambios de idioma se encuentren correctamente identificados.
- **Punto de verificación 6.1:** referido a la organización de los documentos de forma que puedan ser leídos sin hojas de estilo. Por ejemplo, cuando un documento HTML es interpretado sin las hojas de estilo asociadas, debe seguir siendo posible leer el documento.

En la Prioridad 2 se identificaron 87 errores. Uno de ellos se refiere a la invalidez sintáctica de las hojas de estilo (CSS), establecida en el punto de verificación 3.2. La falta de claridad en los enlaces es otro fallo detectado el cual se detecta en el punto de verificación 13.1.

El punto 2.2, relacionado con las combinaciones de los colores de fondo y primer plano, han de tener el suficiente contraste para que sean percibidas por personas con deficiencias de percepción de color o en pantallas en blanco y negro. También hay que verificar que el contraste de color entre fondo y primer plano (en textos e imágenes) resulte suficiente.

En el punto de verificación 3.2, referente a la creación de documentos que estén validados por las gramáticas formales publicadas, hay que verificar que no se utilicen imágenes para transmitir información si existe un lenguaje de marcas apropiado.

En el punto de verificación 5.3, relacionado con la no utilización de tablas para maquetar, a menos que la tabla tenga sentido cuando se alinee, es necesario verificar que el contenido de las tablas tenga sentido cuando se presenta línea por línea. Por otro lado, si la tabla no tiene sentido, se sugiere proporcionar una alternativa equivalente (la cual debe ser una versión alineada).

También se recomienda verificar que no se provoquen destellos en el contenido de la página mediante imágenes, scripts u otros elementos de pro-

Tabla III. Resultados de la evaluación de los portales, mediante HERA y TAW con la WCAG 1.0.

Evaluación			Portales				
			Del Estado Peruano	CONADIS	Servicios al Ciudadano de Perú	Ministerio de la Mujer y Poblaciones Vulnerables	Ministerio de Desarrollo e Inclusión Social
Hera	Prioridad 1	Verificar	13	10	4	8	8
		Bien	1	0	0	0	0
		Mal	0	1	1	0	3
		No aplicable	3	6	12	9	6
	Prioridad 2	Verificar	19	20	17	13	18
		Bien	1	4	9	4	1
		Mal	4	3	4	4	8
		No aplicable	5	2	9	8	2
	Prioridad 3	Verificar	13	11	11	11	13
		Bien	0	3	0	0	2
		Mal	4	1	2	2	2
		No aplicable	2	4	6	6	2
TAW	Prioridad 1	Revisión automática	0	2	22	0	0
		Revisión manual	76	35	157	5	35
	Prioridad 2	Revisión automática	4	3	224	4	5
		Revisión manual	8	22	182	8	7
	Prioridad 3	Revisión automática	0	11	41	1	0
		Revisión manual	9	5	56	9	6

gramación, elemento relacionado con el punto de verificación 7.2.

Por otro lado, en la prioridad 2 en el CONADIS resulta incorrecto la utilización de 15 elementos y 198 atributos de formato HTML para controlar la presentación de elementos que están relacionados con el punto de verificación 3.3. Por otra parte, no se utilizan encabezados para sugerir la estructura del documento (punto de verificación 3.5), y se debió evitar el uso de elementos obsoletos de las tecnologías del W3C, debido a que utilizan 13 elementos y 202 atributos obsoletos en formato de programación HTML, referido al punto 4.1.

También en el punto 3.6, es necesario verificar las cinco listas con viñeta, las cuales no se deben utilizar para conseguir efectos de formato, así como comprobar si existen contenidos de la página que deban presentarse como una lista de elementos, y comprobar los destellos en el contenido de la página mediante imágenes, scripts u otros elemen-

tos de programación referidos al punto 7.2. Por otra parte, es necesario comprobar que los enlaces se encuentran claramente identificados, como establece el punto de verificación 13.1.

Por ejemplo, un elemento positivo que difiere de la evaluación del portal del Estado peruano está relacionado con la no utilización de atributos, ni elementos HTML obsoletos en el portal del CONADIS, pero se detecta como error el cumplimiento del punto 6.2, relacionado con asegurar la actualización de los contenidos dinámicos tras los cambios, lo cual provoca que la página sólo pueda ser utilizada con los scripts activados.

Algunos de los errores encontrados en la prioridad 3 fueron: la falta de información para que los usuarios puedan recibir los documentos según sus preferencias, tales como idioma, tipo de contenido, tamaño de fuente, relacionados con el punto 11.3; la falta de especificación de la expansión de las abreviaturas relacionado con el punto 4.2; así como el punto 9.5

relacionado con proporcionar atajos de teclado y la indicación del lenguaje principal del documento.

La evaluación de los portales del Estado peruano, utilizando la herramienta TAW mediante las WCAG 2.0, también mostró un conjunto de errores en los 3 niveles de prioridad que se muestra en la tabla IV.

En el **Portal del Estado Peruano**, se encontraron en el nivel A, 37 problemas en 7 criterios de éxito, los cuales están relacionados con los puntos de verificación, así como 167 advertencias en 11 criterios. Existen 10 criterios que requieren comprobación manual en su totalidad. Algunos de los errores están relacionados al uso del color, audio y texto, de igual modo que en las WACG 1.0.

En el nivel AA no fue posible verificar los 16 criterios, los cuales requieren revisión completamente manual: perceptibles 4, operable 8, comprensible 4; los 41 problemas encontrados también están relacionados con los componentes de interfaz presentados al usuario.

En el nivel AAA, existen 44 problemas en 10 criterios de éxito, así como 195 advertencias, siendo

necesaria la comprobación manual completa en 27 criterios. Por otra parte, El orden de navegación de los bloques de información mediante tabulación resulta inapropiado, y coloca como uno de los últimos elementos de acceso la barra de menú de navegación. Esto obliga al usuario a que utilice el teclado como dispositivo de acceso, a recorrer todos los objetos que componen la página hasta llegar a dicho menú. El principal aspecto negativo es la dependencia de los elementos dinámicos para la navegación por la página, que dificulta el acceso a las diferentes secciones.

En el **CONADIS**, en el nivel A, son necesarias 9 correcciones en 9 criterios de éxito, como por ejemplo en el punto 1.1.1: en la navegación existen enlaces consecutivos de textos e imágenes para el mismo recurso, en el punto 1.4.4: en la presentación se utilizan tamaños de fuentes absolutos y a excepción de los subtítulos y las imágenes de texto, todo el texto puede ser ajustado sin ayudas técnicas y en el punto 1.4.5: se utilizan imágenes susceptibles de ser sustituidas por texto.

También se encontraron 117 advertencias en 11 criterios de éxito, referidas al punto 1.1: conteni-

Tabla IV. Resultados de la evaluación de los portales del Estado peruano, con WCAG 2.0 y TAW

Portales	Principios	Problemas	Advertencia	Problemas	Advertencia	Problemas	Advertencia
		A		AA		AAA	
Portal del Estado Peruano	Perceptible	21	161	25	163	25	168
	Operable	11	4	11	4	14	4
	Comprensible	3	2	3	6	3	23
	Robusto	2	0	2	0	2	0
Portal CONADIS	Perceptible	6	58	10	76	10	91
	Operable	2	26	2	55	9	55
	Comprensible	0	2	0	6	0	34
	Robusto	1	31	1	32	1	32
Portal de Servicios al Ciudadano	Perceptible	34	158	38	158	38	158
	Operable	20	11	20	11	30	11
	Comprensible	3	2	3	6	3	10
	Robusto	1	21	1	21	1	21
Portal del Ministerio de la Mujer y Poblaciones Vulnerables	Perceptible	1	0	1	0	1	0
	Operable	1	1	1	1	2	1
	Comprensible	1	0	1	0	1	0
	Robusto	5	0	1	0	1	0
Portal del Ministerio de Desarrollo e Inclusión Social	Perceptible	5	0	2	1	10	1
	Operable	6	3	3	1	2	1
	Comprensible	1	0	1	2	1	0
	Robusto	5	0	1	0	1	0

do no textual de imágenes que pueden requerir descripción más larga e imágenes con alt que se encuentran vacías; el punto 1.3, referido a la adaptación de secuencias con significados. En el punto 1.3.2: en la presentación se encuentran mal posicionados algunos aspectos como:

- http://www.conadisperu.gob.pe/templates/bt_arise/css/template.css (con posicionamiento flotante).
- <http://www.conadisperu.gob.pe/plugins/system/jat3/jat3/base-themes/default/css/template.css> (con posicionamiento de forma absoluta).

En el nivel AA, se encontraron 13 problemas en 4 criterios, 169 advertencias en 18 criterios y 16 criterios requieren comprobación manual, como el punto 1.1.1 referido al contenido no textual y el punto 4.4 relacionado con el redimensionamiento del texto en la presentación, donde se utilizan tamaños de fuentes absolutos como por ejemplo:

- http://www.conadisperu.gob.pe/templates/bt_arise/css/template.css
- <http://www.conadisperu.gob.pe/plugins/system/jat3/jat3/base-themes/default/css/template.css>

El nivel AAA, arrojó 20 problemas en 5 criterios y 212 advertencias en 26 criterios y 27 criterios requieren revisión manual. Se mantienen los problemas con los contenidos no textuales (punto 1.1.1), así como el redimensionamiento de textos referido al punto 1.4.4. Por otra parte referente a la navegabilidad del punto 2.4, es necesario verificar el propósito de los enlaces (punto 2.4.4 para el contexto) y para los propios enlaces (punto 2.4.9).

En el nivel A, en el **Portal de Servicios al Ciudadano** se encontraron 58 problemas en 7 criterios de éxito, uno de ellos relacionado con texto e imágenes, también encontrado en los portales anteriores. Igualmente se encontraron 192 advertencias en 8 criterios; en 9 aspectos no fue posible realizar la validación automática y se recomienda validación manual a sus 9 criterios de éxito, por ejemplo el punto 2.1 relacionado con la accesibilidad mediante teclado sin bloqueos (punto 2.1.2). Por otra parte, en el punto 1.3 referente a la adaptabilidad es necesario revisar de forma manual las características sensoriales, y en el punto 1.4 referido a lo distinguible, el uso de color y el contraste.

En el nivel AA, son necesarias 62 correcciones en 8 criterios de éxito, algunos están relacionados con los aspectos legibles del idioma de la página (punto 3.1.1) así como en los propósitos de los enlaces que permiten la navegabilidad (punto 2.4.4). También se encontraron 196 advertencias

que requieren revisión manual, como la accesibilidad mediante el teclado (punto 2.1.1), las páginas tituladas (punto 2.4.2) y los bloques (punto 2.4.1), así como 15 criterios no verificables también requieren revisión completa manual como son: el idioma legible de la página (punto 3.1.1) y de las partes de esta (punto 3.1.2), las características sensoriales de la adaptabilidad (punto 1.3.3) y todos los puntos relacionados con aspectos distinguibles (punto 1.4).

El nivel AAA, presentó un total de 72 problemas en la validación automática en 11 criterios de éxito, manteniéndose los problemas relacionados con el contenido no textual (punto 1.1.1), con la información y relaciones de la adaptabilidad (punto 1.3.1), así como el redimensionamiento de texto (punto 1.3.1). En cuanto a advertencias, 72 problemas en 13 criterios de éxito, como por ejemplo en la accesibilidad mediante el teclado (punto 2.1.1), en la navegabilidad, hay que evitar páginas tituladas (punto 2.4.2) y los bloques (punto 2.4.1), así como en la introducción de datos asistidos (punto 3.3) en todos sus puntos.

En este mismo nivel requieren comprobación manual en su totalidad, 26 criterios, como son: las características sensoriales (punto 1.3.3), todos los puntos relacionados con los elementos distinguibles (punto 1.4) como color, control de audio, imágenes de texto, contraste, sonido de fondo, presentación visual.

La evaluación del **Portal del Ministerio de la Mujer y Poblaciones Vulnerables**, en el nivel A se encontraron 8 problemas, en 5 criterios de éxitos, uno de ellos relacionado con texto e imágenes, encontrados en los portales anteriores, así como la información y los componentes de la interfaz de usuario deben ser presentados a estos, de modo que puedan percibirlos específicamente con la estructura y semántica, al no definir encabezados de primer nivel (elemento h1), y los aspectos referidos a la declaración del idioma en el documento (punto 3.1.1).

Como nuevo aparecen los marcos sin título (punto 4.1.2). También se sugiere que los contenidos implementados mediante el uso de lenguajes de marcas, tengan las etiquetas de apertura y cierre completas y estén anidados de acuerdo a sus especificaciones y sin duplicados (punto 4.1.1). En los aspectos de operatividad en el punto 2.4.2, se encontró como error, que las páginas que contiene el portal se encuentran sin título, por lo que es necesario proporcionar un título para cada página.

Como aspecto favorable en este nivel A se puede destacar que se obtuvo una sola advertencia en 1 criterio, referente al punto 2.4.1, referente a los encabezados de primer nivel.

En el nivel AA, son necesarias 4 correcciones, en 4 criterios de éxitos, lo cual consideramos satisfactorio, si se compara con los resultados obtenidos de las WACG 1.0. Las correcciones a realizar están relacionadas con el punto 1.3 referente a la adaptabilidad, donde es necesario que la información, estructura y relaciones a través de la presentación, puedan ser determinadas de forma programática, o están disponibles de forma textual, se mantiene el aspecto del punto 2.4.2, referente a los títulos de las paginas, el punto 3.1.1 referido al idioma, y los aspectos de nombre, función y valor del punto 4.1.2. Las advertencias que requieren revisión manual, están relacionadas con la necesidad de evitar bloques, debido a que existe un mecanismo para eliminarlos entre las diferentes páginas (punto 2.4.1).

El nivel AAA, presentó un total de 5 problemas en cinco criterios de éxitos, manteniéndose los problemas relacionados con la inexistencia de encabezado de primer nivel (punto 1.3.1), el título de las páginas (punto 2.4.2), el idioma del documento (punto 3.1.1), y los marcos sin títulos (punto 4.1.2). En cuanto a advertencias se mantiene solo un problema, en 1 criterio de éxito, referida al punto 2.4.1, sugiere evitar bloques.

En este mismo nivel se requiere de comprobación manual en su totalidad de 5 criterios, como por ejemplo: las características sensoriales (punto 1.3.3), así como algunos puntos relacionados con los elementos distinguibles como color (punto 1.4.1), imágenes de texto (punto 1.4.5), contraste (punto 1.4.6), presentación visual (punto 1.4.8).

Los resultados del análisis de la accesibilidad web de las páginas interiores de esos portales se muestran en la tabla V.

Los porcentajes de éxito obtenidos muestran como mejor portal en el cumplimiento de los criterios de accesibilidad web al del Ministerio de la Mujer y Poblaciones Vulnerables con 57,14 % seguido del portal del Ministerio de Desarrollo e Inclusión Social y, aunque el portal de Servicios al Ciudadano no es el que obtiene un menor porcentaje, su resultado no es alentador, a su vez es el portal que mayor cantidad de formularios tablas posee, necesarias para el uso de los usuarios del sitio.

7.2. Resultados de la evaluación de la Fase 2

La evaluación realizada con el AIS al portal de Estado peruano, arrojó que existen 5 errores HTML como: algunos tipos de documentos no permiten hacer link, o enlaces que no aparecen y 2 advertencias, relacionadas con la no generación de identificador de sistemas y algunas referencias no terminadas. Sin embargo la validación CSS, no mostró errores.

En cuanto a la navegación accesible (aspecto 2), se encuentran enlaces, mediante los cuales es imposible realizarla:

- <http://www.presidencia.gob.pe/index.asp>
- http://www.peru.gob.pe/areas/pep_areas_galeria.asp
- <http://www.facebook.com/GobiernoElectronicoONGEI>

Tabla V. Resultados de la evaluación realizada a los Portales del Estado evaluando 2 páginas web

Portal	Comprobaciones				Porcentaje de éxito
	Bien	Mal	Pocos fallos	No aplicable	
Portal del Estado Peruano	6	20	4	1	29,63
Portal del Consejo Nacional para la Integración de la Persona con Discapacidad	10	18	3	2	44,23
Portal Servicios al Ciudadano	5	21	5	4	31,25
Portal del Ministerio de la Mujer y Poblaciones Vulnerables	14	15	4	0	57,14
Ministerio de Desarrollo e Inclusión Social	11	13	6	2	53,85

Además, algunos encabezados mantienen la misma dificultad encontrada en la evaluación automática, referente a que los encabezados no utilizan de forma correcta, el nivel de profundidad. Respecto a las imágenes, en la página principal, se detectó el mismo problema de la revisión automática, referido al contraste de colores, asimismo, se detectó que estas no ofrecen alternativas mediante textos, y en algunos casos, la alternativa fue concebida sólo mostrando el nombre del fichero de la imagen.

La estructura de las tablas de datos, que aparecen en la página se encuentran maquetadas de forma aceptada, no obstante la W3C aconseja el uso del elemento "summary", que permite agregar resúmenes a cada una de las tablas, lo mismo que usar unos atributos como el "caption", para realizar los pie de páginas, así como agregar siempre encabezados a filas y columnas, a través del elemento "th", que identifique adecuadamente, las filas y columnas, y de esta forma facilitar su lectura por los lectores de pantalla.

De la evaluación realizada con el AIS a la página del CONADIS, se obtuvo: que la validación HTML, encontró solo 2 errores, uno relacionado con el tamaño de la fuente y el otro referido a la no atribución de determinados elementos a los *click* "data-mce-bogus". La evaluación del CSS, arrojó 56 errores, relacionados casi todos con un conjunto de enlaces, que no pueden ser revisados por la herramienta y existe una advertencia sobre el constante redireccionamiento del enlace <http://twitter.com/conadisperu>, redireccionado a <https://twitter.com/conadisperu>.

Los dos errores relacionados con la utilización de HTML, fueron:

- No se encuentran atributos relacionados con "onClick", para el cual se sugiere incrementar el tamaño de la fuente.
- No se atribuyen elementos relacionados con "data-mce-bogus", para lo cual se sugiere que los desarrolladores de la página usen la siguiente sintaxis `xt-align: center;"><br data-mce-bogus="1" /><a href="/web/documentos/2013/Conv.`

La validación CSS, encontró los siguientes errores:

- (* + html .ja-typo-bubblelect), este resultado indica que la propiedad zoom no existe.
- (* + html .ja-typo-legend), este resultado también indica la existencia de errores en la propiedad zoom.

Así como un total de 88 advertencias en varias líneas de programación de la página. En cuanto a

la accesibilidad se encontraron un total de 56 enlaces, por los cuales no es posible realizar la navegación, algunos de ellos son:

- <http://www.conadisperu.gob.pe/media/system/js/core.js>
- http://www.conadisperu.gob.pe/images/revista8_portada.jpg
- <http://www.conadisperu.gob.pe/media/system/css/modal.css>

Por otra parte, pocas imágenes tienen contenido textual, y aunque algunos criterios de contraste de color han sido relevantes, la W3C, sugiere que debe existir un contraste bien definido entre imágenes y texto. Una de las imágenes hace uso del color como ayuda semántica, para transmitir información sobre el contenido de éstas; aunque esto no supone una deficiencia en cuanto a accesibilidad web, es desfavorable para personas con discapacidad cognitivas, las cuales necesitarán guías para poder tener una navegación accesible.

La validación manual realizada con el AIS al portal de servicios al ciudadano, relacionada con la utilización de HTML, presentó un total de 25 errores, algunos de los cuales fueron detectados con la evaluación automática, los cuales son:

- Se requiere atributo "type" no especificado
- Algunos documentos donde no se encuentran los elementos "script"; por lo que son asumidos o encontrados como "body" start-tag
- Existen atributos definidos como "Title", que no lo son
- Se requieren atributos "alt" los cuales no han sido especificados (siendo este error el de mayor aparición entre los 25)
- El final de etiquetas para "body" omitido, pero OMITTAG no ha sido especificado

El validador de utilización de hojas de estilo CSS, encontró 21 errores, algunos relacionados con el color. Sin embargo, solo se encontraron siete enlaces que dificultan la navegación accesible, errores que están relacionados con problemas en el servidor interno donde se encuentra el enlace o accesos a 'javascript' URIs que han sido desactivados. Ellos son:

- http://www.pcm.gob.pe/Transparencia/Resol_ministeriales/2010/DS-064-2010-PCM.pdf
- `javascript:%20Abrir_ayuda('ayuda/ayuda.htm')`
- `javascript:Enlazar2('bus/info/Vermas.html')`
- <http://www.empresas.gob.pe/>

- <http://www.notariadigital.org.pe/ol-it-portal/solicitud.do>

Las imágenes, siguen presentando dificultad en cuanto a su contraste, color y sin contenido textual. En algunas tablas que recopilan información, sus columnas no presentan encabezado, lo cual puede dificultar la lectura de un lector de pantalla.

En la evaluación realizada con el AIS al portal de del Ministerio de la Mujer y Poblaciones Vulnerables se encontraron, cuatro errores en la utilización de HTML:

- No se encuentran documentos "type".
- Se require atributos "type", los cuales no han sido especificados.
- Se encuentran extraviados los requisitos del sub-elemento a HEAD

Una de sus mayores deficiencias encontradas también en la evaluación con Hera y TAW, es la relacionada con las hojas de estilo, cuya validación del CSS arrojó un total de 180 errores y 95 advertencias, relacionadas con propiedades no definidas para el color y errores de análisis sintáctico, siendo esta última la de mayor incidencia.

Se mantienen los problemas detectados con las imágenes en la evaluación automática, relacionada con que estas no brindan información textual, y como texto alternativo, solo muestran el nombre de la imagen.

Por último, en el Portal del Ministerio de Desarrollo e Inclusión Social, con AIS, referida a la utilización del HTML, se detectaron 13 errores y una advertencia, algunos de los cuales se muestran a continuación:

- *Document type does not allow element "style" here*
- *Document type does not allow element "link" here*
- *Character "<" is the first character of a delimiter but occurred as data*
- *Document type does not allow element "div" here; assuming missing "li" start-tag*
- *Start tag was here*

La utilización del CSS, arrojó un total de 96 errores y 45 advertencias, relacionadas con el color y errores de análisis sintáctico, siendo esta última la de mayor incidencia, y se mantienen los problemas detectados con las imágenes en la evaluación automática, relacionada con que no brindan información textual, y como texto alternativo, solo muestran el nombre de la imagen.

Se encuentran 39 enlaces en los que es imposible realizar un análisis de accesibilidad, algunos porque no se encuentra la página, o porque es restringido el acceso, o porque existe problemas con el 'javascript' URIs, como por ejemplo:

- http://www.midis.gob.pe/index.php/es/ico_doc.gif
- <http://www.youtube.com/midisperu>
- <http://www.contraloria.gob.pe/>
- http://www.midis.gob.pe/images/banners_laterales/tupa1.jpg
- <https://www.facebook.com/MidisPeru>

También algunas tablas que recopilan información, sus columnas y filas no presentan encabezado, lo cual puede dificultar la lectura de estas a través de un lector de pantalla.

8. CONCLUSIONES FINALES

El principal objetivo del presente trabajo ha sido evaluar la accesibilidad de las páginas principales de los portales del Estado en Perú de acuerdo con la norma WCAG 1.0, y de la Resolución Ministerial 126-2009-PCM de la Presidencia del Consejo de Ministros donde establece para las instituciones públicas del Estado peruano, la obligatoriedad del cumplimiento de las Prioridades 1 y 2. Los resultados arrojaron que los portales analizados contienen diferentes errores que dificultan la accesibilidad para personas con discapacidad.

El análisis de los puntos de verificación por prioridad en los portales evaluados arrojó deficiencias en materia de accesibilidad web en las 3 prioridades. El único portal que no presentó elementos y atributos obsoletos en el formato de programación HTML, fue el del CONADIS.

La evaluación de la accesibilidad Web de los principales portales del Estado peruano, a partir de las pautas de accesibilidad para el contenido Web 2.0, con el validador TAW, arrojó un conjunto de deficiencias en uno de los tres niveles analizados como: errores relacionados al uso del color, audio y texto, contenido no textual para imágenes que pueden requerir descripción más larga e imágenes con alt, que se encuentran vacías, errores con los componentes de interfaz presentados al usuario, enlaces sin contenidos y el orden de navegación de los bloques de información mediante tabulación resulta inapropiado.

Los resultados obtenidos en la evaluación anterior a la página de inicio, no fueron muy diferentes al analizar los criterios de éxitos, de las páginas de interiores de los portales.

La accesibilidad web de los portales del Estado peruano, utilizando la herramienta AIS Web Accessibility Toolbar, arrojó en los aspectos analizados, lo siguiente:

- **La utilización de HTML y CSS:** La evaluación HTML detectó varios problemas con la no determinación de elementos, *Type*, así como con la no generación de identificador de sistemas y algunas referencias no terminadas. La evaluación del CSS, determinó un conjunto de errores, relacionados en su gran mayoría con los enlaces, que no pueden ser revisados por la herramienta y por el constante redireccionamiento de algunos enlaces. Solo el portal del Estado peruano no presentó deficiencias en la utilización de las hojas de estilo CSS.
- **La navegación accesible:** Un conjunto de enlaces por los cuales no se puede realizar la navegación y los encabezados no utilizan de forma correcta, el nivel de profundidad.
- **Descripción de imágenes:** En todos los portales las imágenes no ofrecen alternativas mediante textos, y en algunos casos, la alternativa fue concebida, solo mostrando el nombre del fichero de la imagen y su contraste con el texto es deficiente.

- **La estructura de las tablas de datos:** Algunas tablas se encuentran maquetadas de forma aceptable, principalmente en el portal del Estado peruano. No obstante es necesario agregar encabezados a filas y columnas en los casos necesarios, a través del elemento "th", que identifique adecuadamente, las filas y columnas, y así facilitar su lectura por los lectores de pantalla.

Es necesario que los desarrolladores de los portales de las correspondientes instituciones del Estado peruano analicen, el conocimiento e implementación de las pautas de accesibilidad de la WCAG 2.0, en el marco legislativo sobre accesibilidad web revisar lo referente a la viabilidad de mantener el cumplimiento de la WCAG 1.0, sustituyendo esta por la WCAG 2.0, en función de los nuevos proyectos que se gestan en el país, para tomar acciones favorables en cuanto al diseño de portales accesibles, de manera que todos los usuarios accedan a la información, independiente del equipo de acceso o de la tecnología utilizada. Así como definir el Modelo de Accesibilidad Web, que guiará al Estado peruano, no definido, en el marco legislativo, ni en las proyecciones para los 5 años posteriores.

9. REFERENCIAS

- Acevedo, J.J.; Mariño, S.I.; Godoy, M.V. (2011). Análisis accesibilidad de los sitios web de noticias de las provincias del Nordeste Argentino. *No solo usabilidad*, 10. Recuperado de <http://nosolousabilidad.com>
- Caballero-Cortés, L.; Faba-Pérez, C.; Moya-Anegón, F. (2008). Evaluación comparativa de la accesibilidad de los espacios web de las bibliotecas universitarias españolas y norteamericanas. *Investigación Bibliotecológica*, 23(47), 45-66. Recuperado de <http://www.ejournal.unam.mx/ibi/vol23-47/IBI002304703.pdf> [consulta: 12/03/2014].
- Chacón-Medina, A.; Chacón-López, H.; López-Justicia, M. D.; Fernández-Jiménez, C. (2013). Dificultades en la Accesibilidad Web de las Universidades Españolas de acuerdo a la Norma WCAG 2.0. *Revista Española de Documentación Científica*, 36 (4), 1-13. Recuperado de <http://dx.doi.org/10.3989/redc.2013.4.1009> [consulta: 28/05/2014].
- Córdova Solís, Miguel Ángel. (2012). Estudio comparativo de accesibilidad web en portales informativos de universidades peruanas de educación a distancia. *Calidad y Accesibilidad de la Formación Virtual. Alcalá de Henares: Universidad de Alcalá*, 63-73.
- Díaz, F.; Harari, I. y Amadeo, P. (2008). *Evaluación de Accesibilidad de Sitios Web oficiales: Ministerios y Municipios de la Provincia de Buenos Aires*. La plata (Argentina): Universidad Nacional de La Plata. Laboratorio de investigación en nuevas tecnologías informáticas. Recuperado de http://www.linti.unlp.edu.ar/uploads/docs/proceso_de_evaluacion_de_accesibilidad_de_sitios_web_oficiales.pdf [consulta: 16/02/2014].
- Díez Canseco, J. (2009). Situación de la discapacidad en la Región Andina (Legislación y Políticas de Estado): Organismo Andino de Salud – Convenio Hipólito Unanue (ORAS-CONHU).
- Hassan Montero, Y.; Martínez Fernández, F.J. (2003). Guía de evaluación heurística de sitios web. *No Sólo Usabilidad e-magazine*. <http://www.nosolousabilidad.com/articulos/heuristica.htm> [consulta: 14/04/2014].
- Hernández, R.; Amado-Salvatierra, H.R. (2012). Evaluación de la accesibilidad de portales web en instituciones educativas en el área de Centroamérica. *En Calidad y Accesibilidad de la Formación Virtual. Alcalá de Henares: Universidad de Alcalá*, 74-80.
- Hilera, J.R.; Fernández, L.; Suárez, E.; Vilar, E.T. (2013). Evaluación de la accesibilidad de páginas web de universidades españolas y extranjeras incluidas en rankings universitarios internacionales. *Revista Española de Documentación Científica*, 36 (1), 1-16. Recuperado de <http://dx.doi.org/10.3989/redc.2013.1.913> [consulta: 25/03/2014].

- Huerta Peralta, J. (2006). Discapacidad y Accesibilidad: La dimensión desconocida. In F. E. d. C. d. Perú (Ed.), (pp. 189). Lima, Perú: Comisión Especial de Estudio sobre Discapacidad del Congreso de la República.
- INTECO (2008). *Estudio-Diagnóstico de accesibilidad y calidad web en la pyme española* [en línea]. Madrid: Instituto Nacional de Tecnologías de la Comunicación. https://www.incibe.es/file/IFrLER_WNidc6KV2OP1O1Q [consulta: 20/04/2014].
- ISO/IEC 40500. (2012). Information technology W3C Web Content Accessibility Guidelines (WCAG) 2.0.
- Karhu, M.; Hilera, J.R.; Fernández, L.; Rios, R. (2012). Accessibility and readability of university websites in Finland. *En Calidad y Accesibilidad de la Formación Virtual, Alcalá de Henares: Universidad de Alcalá*, 151-158.
- Medina, J. (2012). *Evaluación de la sede web del Centro Regional de Información sobre Desastres (CRID): aplicación del cuestionario BRACAD*. Tesis para optar el grado de Magíster, Universidad de Salamanca, Salamanca, España.
- Montano, J. (2014). *Evaluación de la accesibilidad y usabilidad del sitio web de la Unión Cuba-Petróleo. Tesis de Diplomado en Gestión de la Información*, Centro de Desarrollo Profesional en Ciencias de la Información (PROINFO), La Habana, Cuba.
- Naftali, M.R. (2010). *Análisis e Integración de métricas para la Accesibilidad Web*. (Tesis de grado), Universidad de Buenos Aires, Argentina.
- ONU. (2008). Convención sobre los Derechos de las Personas con Discapacidad.
- PNUD. (2009). Compendio de Legislación sobre discapacidad: Marco Internacional, Interamericano y de América Latina (pp. 512): Programa de las Naciones Unidas para el Desarrollo.
- Quiliano, R.; Ríos, A. (2012). Estudio de accesibilidad web en las bibliotecas virtuales en salud en Perú. Ponencia presentada en el 9º Congreso Regional de Información de Ciencia y Salud, Crics, OPS, Octubre 22-24, Washington, D.C. Recuperado de: http://www.crics9.org/es/files/2012/12/Id_17_PE_ESTUDIO1.pdf [consulta: 21/04/2014].
- Rosales, P. (2011). *Análisis Crítico descriptivo de la normativa de la discapacidad en Argentina con énfasis en el sistema de salud y en el marco de la convención sobre los derechos de las personas con discapacidad - Ley N° 26.738*. Tesis de Master, Universidad Uisalud, Buenos Aires, Argentina.
- Segovia, C. (2007). Accesibilidad e Internet... para que todas las personas con distintas capacidades y recursos puedan acceder a Internet. Madrid: SIDAR. Recuperado de <http://sid.usal.es/idos/F8/FDO19018/accesibilidadeinternet.pdf> [consulta: 23/02/2014].
- Serrano Mascaraque, E.; Moratilla Ocaña, A.; Olmeda Martos, I. (2010). Métrica para la evaluación de la accesibilidad en Internet: propuesta y testeo. *Revista Española de Documentación Científica*, 33(3), 378-396. <http://dx.doi.org/10.3989/redc.2010.3.719>
- Varela, C.; Miñán, A.; Hilera, J.R.; Restrepo, F. A.; Amado, H.; Córdova, M.A.; y Villaverde, A. (2012). *Estándares y legislación sobre accesibilidad web*. Paper presented at the Actas del IV Congreso Internacional ATICA 2012 Loja, Ecuador.
- W3C. (1999). Web Content Accessibility Guidelines (WCAG) 1.0. World Wide Web Consortium. <http://www.w3.org/TR/WCAG10/> [consulta: 16/02/2014].
- W3C. (2011). Guía breve de tecnologías W3C. España: World Wide Web Consortium (W3C). Recuperado de <http://w3c.es/Divulgacion/GuiasBreves> [consulta: 16/02/2014].
- Zubillaga del Río, Ainara (2010). *La Accesibilidad como elemento del proceso educativo: Análisis del Modelo de Accesibilidad de la Universidad Complutense de Madrid para atender las necesidades de los estudiantes con discapacidad*. Tesis doctoral. Universidad Complutense de Madrid, España. (T32369)



ESTUDIOS / RESEARCH STUDIES

Aplicación del modelo de datos RDF en las colecciones digitales de bibliotecas, archivos y museos de España

Andreu Sulé*, Miquel Centelles*, Jorge Franganillo*, Jesús Gascón*

*Universitat de Barcelona. Departament de Biblioteconomia i Documentació
Correos-e: sule@ub.edu, miquel.centelles@ub.edu, franganillo@ub.edu, gascon@ub.edu

Recibido: 25-02-2015; 2ª versión: 30-06-2015; Aceptado: 09-07-2015.

Cómo citar este artículo/Citation: Sulé, A.; Centelles, M.; Franganillo, J. y Gascón, J. (2016). Aplicación del modelo de datos RDF en las colecciones digitales de bibliotecas, archivos y museos de España. *Revista Española de Documentación Científica*, 39(1): e121. doi: <http://dx.doi.org/10.3989/redc.2016.1.1268>

Resumen: Se analiza cómo y en qué grado se aplica el modelo RDF en las principales colecciones digitales españolas de materiales patrimoniales. Se introduce este modelo y también las iniciativas de Datos Abiertos y Datos Enlazados. A continuación se examinan 51 repositorios digitales y de cada uno de ellos se determina si expresan sus registros en RDF, ofrecen un punto de consulta SPARQL consultable por agentes externos y si usan referencias como valor de las propiedades. A partir de los resultados se describen los modelos EDM de *Europeana* y *OntoWeb*. Se concluye que la aplicación de RDF es desigual, demasiado condicionada al uso de aplicaciones que convierten automáticamente los registros en tripletes RDF, que se dan pocos puntos SPARQL para la consulta externa y que el empleo de referencias se vincula al uso de aplicaciones con diferentes modelos: EDM u *OntoWeb*. Se recomienda que las colecciones han de enriquecer sus datos y determinar niveles de agregación de los datos RDF generados para difundirse, hacerse accesibles y adaptarse a la web semántica.

Palabras clave: RDF; modelos de datos; repositorios patrimoniales; consulta SPARQL; referencias; URI; *Europeana Data Model*; *OntoWeb*; datos abiertos; datos enlazados.

Implementation of the RDF data model in digital collections of Spanish libraries, archives and museums

Abstract: The article discusses how and to what extent the RDF data model is applied in major Spanish digital collections of heritage materials. This model, as well as Open Data and Linked Data initiatives, are introduced. Fifty-one digital repositories were analysed to determine whether they expressed their records in RDF, offered SPARQL query points searchable by external agents, and used references as property values. The *Europeana EDM* and *OntoWeb* models are also described. It is concluded that the use of RDF is unequal and excessively conditioned by the use of applications that automatically convert records into RDF triples. Few of the collections analysed give SPARQL points for external queries. Also, the use of references is linked to applications using different models: EDM or *OntoWeb*. Collections should enrich their data and define aggregation levels for generated RDF data in order to be disseminated, made accessible, and adapted to the semantic web.

Keywords: RDF; data models; heritage repositories; SPARQL query; references; URI; *Europeana Data Model*; *OntoWeb*; open data; linked data.

Copyright: © 2016 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0.

1. INTRODUCCIÓN

1.1. La web semántica y el modelo de datos RDF

Las bibliotecas, archivos y museos y, más concretamente, sus depósitos digitales, atesoran generalmente ingentes cantidades de datos (organizados en forma de documentos) y metadatos (organizados en registros bibliográficos y de autoridad), que presentan altos niveles de estructuración, de tratamiento desde el punto de vista semántico y de interoperabilidad. Más allá de objetivos vinculados a la conservación y preservación de sus fondos, estos tres tipos de instituciones necesitan darlos a conocer y hacerlos más accesibles a otras instituciones similares, pero sobre todo a sus usuarios, sean particulares, instituciones o empresas. Todos estos factores resultan especialmente interesantes para la web semántica, y, más específicamente, la web de datos y su propósito de hacer los contenidos más accesibles (leíbles y comprensibles) para las máquinas.

La web semántica pone a disposición de bibliotecas, archivos y museos un modelo de datos para la representación de información en la Web, el Resource Description Framework, RDF¹, y unas tecnologías que, en conjunto, permiten dotar de significado a los datos y a las transacciones de datos en la Web. Todos estos componentes han sido estandarizados en forma de normas, especificaciones, recomendaciones, etc. por el W3C².

Una declaración RDF constituye la forma más simple de expresión de un metadato, mediante grafos de tripletas compuestas por un sujeto –la cosa, el recurso, sobre el que se declara algo–, un predicado –el aspecto que se describe– y un objeto –el valor de lo que se describe sobre el sujeto–. Los recursos que constituyen el sujeto y el predicado tienen que ser designados por identificadores uniformes de recursos, URI³, mientras que el recurso que se sitúa en el objeto de la declaración puede ser expresado mediante URI o literales⁴. Además, los recursos de los tres componentes de las declaraciones pueden vincularse a ontologías, conceptualizaciones explícitas de ámbitos concretos de la realidad, expresados por lenguajes normalizados –RDF Schema, SKOS, OWL⁵, y formados, como componentes básicos, por conceptos y relaciones entre conceptos. Al vincular los recursos situados en sujetos y objetos de declaraciones RDF a los conceptos y los predicados a las relaciones, el significado de todos ellos puede elevarse a un marco de representación más amplio que la expresión concreta de una declaración, puede ser compartido por diferentes productores y consumidores de datos, y puede ser sometido a restricciones se-

mánticas expresadas formalmente, muy útiles en el momento que son procesadas por los sistemas automáticos.

Por otro lado, la estructura de grafo permite el procesamiento de la información existente independientemente de que sea completa o detallada o del grado de estructuración de sus fuentes. Al ser una estructura simple, no está sujeta a los cambios de esquema o de formato habituales en las actualizaciones de bases de datos relacionales, y permite el fácil intercambio de datos entre sistemas o versiones diferentes, sin pérdida de datos.

1.2. Comparativa con otros modelos/sintaxis que compiten con RDF

Existen diferentes formatos, o sintaxis, que permiten la codificación de las tripletas que constituyen un grafo RDF y su publicación como datos enlazados. Son los llamados formatos de serialización de RDF, algunos de los cuales han sido normalizados por el W3C.

- Sintaxis XML para RDF, o, simplemente, RDF/XML⁶.
- Terse RDF Triple Language, abreviado Turtle, que permite expresar un grafo RDF en una forma textual compacta. TriG es una extensión de Turtle para representar conjuntos de datos RDF completos⁷.
- RDFa que permite incorporar datos estructurados y semánticamente enriquecidos de acuerdo con el modelo RDF en código HTML y XHTML, mediante un conjunto de atributos⁸.

A priori, la elección de un patrón específico de publicación de datos enlazados determina la preferencia por un formato de serialización específico. Así, en los casos en que la provisión de los datos se realiza incrustados en ficheros HTML, RDFa es la solución idónea. Para el resto de patrones de publicación, entre los que podemos situar la descarga de ficheros (*dumping*), la generación de grafos desde bases de datos relacionales, o desde almacenes de tripletas, etc., rivalizan RDF/XML, Turtle, y otros⁹. Turtle ha tenido éxito por la facilidad de escritura y lectura para las personas. No obstante, cada sistema aplica, de origen, un formato “nativo” específico. En cualquier caso, es importante tener en cuenta la existencia de mecanismos de mapeo y transformación entre formatos.

Más allá de los formatos de serialización, se han generado otros modelos de datos diferentes a RDF, que permiten representar información estructurada y semánticamente rica sobre los datos. Entre los que mayor éxito han alcanzado encontramos

los Microdatos y JSON-LD, ambos impulsados y normalizados por la W3C. A priori, ambos pueden contemplarse como competidores del modelo RDF en la consecución de la web de datos semánticamente enriquecidos.

Los Microdatos¹⁰ son grupos de pares nombre-valor —conocidos como ítems—, que se incorporan en elementos HTML5 como atributos, para describir a qué tipo de entidad hace referencia el contenido de dichos elementos, y qué propiedades se describen sobre dichas entidades. Los valores de las propiedades son, mayoritariamente, segmentos del propio contenido. Existen múltiples vocabularios para la asignación de las entidades y propiedades de las entidades; entre los más conocidos y utilizados se encuentra Schema.org, impulsado por los grandes buscadores generalistas de la web. La facilidad de aplicación de esta tecnología, que puede incorporarse en el flujo de trabajo general de la edición de contenidos en la web, y su potencialidad en la mejora de los resultados de las búsquedas en la web y del posicionamiento de los contenidos, han sido importantes factores de impulso. No obstante, en el momento actual, estas potencialidades de los Microdatos vinculadas a la búsqueda semántica y al posicionamiento web no se han desarrollado completamente. Por otro lado, los Microdatos no son legibles, en origen, por analizadores de datos enlazados y si bien se han desarrollado herramientas para la extracción de Microdatos desde HTML y su reconversión al modelo RDF¹¹, el flujo de trabajo en un contexto real puede complicarse de forma considerable.

JSON-LD¹² es una alternativa a los formatos de serialización de RDF que está ganando terreno de forma rápida en la publicación de datos enlazados. Esto se debe a su doble naturaleza. Por un lado, está basado en el modelo de datos JSON, lo que le confiere la capacidad de modelar los datos en forma de texto comprensible por humanos (pares atributo-valor), para transmitir datos entre servidores y aplicaciones web, de forma independiente a cualquier lenguaje de programación específico. Por otro lado, su modelo de datos es, en gran medida, compatible con RDF; las limitaciones del original JSON han sido superadas proporcionando apoyo a la referencia a recursos mediante el uso de URI, y proporcionando un mecanismo de vinculación de términos de vocabularios a sus URI correspondientes¹³. La diferencia fundamental se encuentra en el hecho que JSON-LD admite tripletes con sujetos, predicados y objetos expresados por URI, nodos vacíos y literales en cualquiera de las tres posiciones¹⁴. JSON-LD está sirviendo como pasarela para la incorporación de datos enlazados en

contextos críticos para su desarrollo y éxito; como son entornos JavaScript, servicios Web y bases de datos NoSQL, sin que se ponga en peligro la continuidad de RDF, ya que, siguiendo la tendencia de otros formatos y modelos de datos, se han desarrollado especificaciones para la interpretación de JSON-LD como RDF, y para la serialización de RDF como JSON-LD.

Podemos concluir que la selección de modelos de datos diferentes, aunque muy próximos, a RDF, como son Microdatos y JSON-LD, para el desarrollo de aplicaciones de datos enlazados está profundamente vinculada a preferencias de contextos y comunidades de uso. Lo que pareciera, a priori, una amenaza puede verse como una oportunidad de expansión, apoyada por la posibilidad de transformación entre todos los modelos de datos y formatos de serialización.

1.3. RDF, Open Data y Linked Data

Para la publicación —y, por lo tanto, también el consumo— de datos RDF resultan fundamentales dos iniciativas complementarias: Datos Abiertos (Open Data) y Datos Enlazados (Linked Data)¹⁵. La primera iniciativa no prescribe, *per se*, la adopción del modelo de datos RDF. Se trata de una apuesta por la interoperabilidad de los datos desde una perspectiva legal. De acuerdo con la definición de Open Bibliographic Data¹⁶, los Datos Abiertos son, en esencia, usables, reusables y redistribuibles libremente y se manifiestan mediante la aplicación de licencias y formatos abiertos. Esta iniciativa ha tenido un impacto decisivo en el ámbito de la administración pública, cuyo marco jurídico fue definido por primera vez a escala europea con la *Directiva 2003/98/CE*¹⁷, y en España, con su transposición al derecho propio en la *Ley 37/2007*¹⁸. Es importante tener en cuenta que en estas primeras disposiciones, los datos y metadatos de bibliotecas, archivos y museos quedaban explícitamente excluidos del amparo de la regulación. No obstante, la revisión de aquella primera directiva mediante la *Directiva 2013/37/UE*¹⁹ ha enmendado esta situación al considerar los beneficios que pueden aportar los datos generados en esas instituciones, no sólo desde un punto de vista cultural y educativo, sino también en oportunidades de negocio. Esta directiva debe transponerse al derecho propio de España, como muy tarde el 18 de julio de 2015, y, en el momento de elaborar este artículo, el *Proyecto de Ley por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público* ha sido publicado en el *BOCG*²⁰ y se ha encomendado su aprobación por el organismo legislativo.

La modificación del marco normativo de los Datos Abiertos en España va a tener implicaciones relevantes en relación con el desarrollo de esta iniciativa para las Administraciones y organismos del sector público, en general, y, especialmente, para las bibliotecas, archivos y museos.

Por un lado, los nuevos textos normativos abren explícitamente el marco de actuación de los Datos Abiertos tanto a los documentos —o datos— como los metadatos que estos tienen vinculados. En lo que se refiere a los formatos, establecen que, siempre que sea posible, tanto los datos como los metadatos deben ser proporcionados en formatos abiertos y legibles por máquina, y deben cumplir estándares y normas abiertas. Además, los metadatos deben ofrecerse con los niveles más elevados de precisión y desagregación.

Por otro lado, es importante tener en cuenta que una motivación fundamental para la incorporación de estas instituciones al marco normativo de los Datos Abiertos es la constatación de que sus fondos, y, especialmente, los que resultan de proyectos de digitalización, son clave para el desarrollo de productos y servicios en diferentes sectores empresariales (se citan, concretamente, el aprendizaje y el turismo). De hecho, el proyecto de ley española insta a las Administraciones Públicas a velar por la reutilización de los datos para fines tanto comerciales como no comerciales (artículo 4.2), e incorpora medidas de impulso y flexibilización del régimen administrativo de la reutilización de los datos y metadatos (diversificación de las modalidades de reutilización, limitación de las condiciones incorporadas en las licencias, etc.)

Especialmente relevante, en relación con el impulso de la reutilización de datos y metadatos, es el requerimiento a las instituciones a adoptar una actitud proactiva en lo que se refiere a los mecanismos de búsqueda y recuperación de los datos por parte de los usuarios finales. En concreto, el artículo 4.5 del proyecto de ley española prevé medidas que comprometen a todas las Administraciones y organismos del sector público: la creación de sistemas de gestión documental enlazados con sistemas ofrecidos por otras administraciones que faciliten la recuperación de los datos por parte de los ciudadanos, y la provisión de herramientas de búsqueda de datos abiertos para su reutilización. Adicionalmente, ordena a la Administración General del Estado mantener “un catálogo de información pública reutilizable, correspondiente al menos al ámbito de la Administración General del Estado y a sus organismos públicos vinculados o dependientes.” (*Proyecto de Ley por la que se modifica la Ley 37/2007*, p. 5).

La segunda iniciativa a la que hemos hecho referencia, Datos Enlazados, aborda la perspectiva técnica de la interoperabilidad y es una apuesta clara por el modelo de datos RDF y por otras dos tecnologías de la web semántica: los URI y el lenguaje de búsqueda en datos RDF denominado SPARQL²¹. Ya son clásicos los cuatro principios rectores de este método de exponer, compartir e interconectar datos en la Web enunciados originalmente por Berners-Lee (2006):

1. Deben utilizarse URI para identificar los recursos en cada uno de los componentes de las declaraciones RDF: en el sujeto, en la propiedad y, con prioridad sobre la alternativa de los literales, en el objeto.
2. Deben utilizarse URI creados de acuerdo con el protocolo HTTP, de forma que puedan ser consultados y desreferenciados en la web por las personas y, sobre todo, por los sistemas automáticos.
3. Debe proporcionarse información útil sobre los recursos identificados con los URI desreferenciables, para cuando alguien los consulte. Con esta finalidad, es necesario utilizar estándares de la web semántica, como RDF y SPARQL.
4. Deben establecerse enlaces con otros recursos (utilizando sus URI) en el momento de publicar datos en la web, de forma que se puedan descubrir más datos.

La publicación de datos enlazados significa la creación de conjuntos de datos (*data sets*) en el contexto de una o más entidades, donde se agregan y mantienen las declaraciones RDF, formando un grafo global, sin los límites de los registros (bibliográficos y de autoridades) y catálogos locales donde se generaron inicialmente. De esta forma, están listos para:

- Enlazarse, en origen, con datos RDF de otros depósitos digitales e incluso de otros conjuntos de datos ajenos a las colecciones documentales.
- Ser enlazados por conjuntos de datos externos que quieran ser enriquecidos.
- Ser consumidos directamente por aplicaciones que los necesiten para la generación de nuevos recursos y servicios, mediante descargas masivas (*dump*) o mediante consultas selectivas de datos con el uso de tecnologías de consulta de datos RDF, ajenas al paradigma de la web semántica (RESTful) o específicamente diseñadas para ésta (puntos de acceso SPARQL).

Los beneficios de la aplicación del modelo de Datos Abiertos Enlazados²² a los datos de institucio-

nes del patrimonio cultural han sido señalados por el W3C Library Linked Data Incubator Group²³. El ahorro de costes en el enriquecimiento de los datos que se ponen a disposición de los usuarios, el incremento de la difusión y accesibilidad de las colecciones y, más concretamente, su (re)utilización en proyectos culturales y educativos y también en iniciativas empresariales. En conjunto, es una buena oportunidad para impulsar la visibilidad de estas colecciones y su posicionamiento en el ámbito de los recursos de información. También es verdad que la adaptación al modelo de datos RDF y la iniciativa de publicación en forma de Datos Abiertos Enlazados imponen unas condiciones técnicas, legales y políticas que será necesario abordar, en especial en el momento en que los datos y metadatos de bibliotecas, archivos y museos deban incorporarse al conjunto de los datos abiertos de la administración del estado.

El modelo se ha venido aplicando y se ha impuesto como paradigma en la comunidad que trabaja en bibliotecas digitales. Como ejemplos, a parte de los de centros españoles que se comentan más adelante, se puede citar el *Zentrale Verzeichnis Digitalisierter Drucke*²⁴, depósito centralizado de 108 colecciones digitales alemanas, mayoritariamente de fondos patrimoniales y antiguos, con todos los metadatos recogidos e integrados en una base de datos única de metadatos para acelerar las búsquedas. Entre estas colecciones destaca la *Digitale Bibliothek* del Münchener Digitalisierungszentrum (Bayerischen Staatsbibliothek)²⁵, con cerca de 907.500 títulos digitalizados. Ambas trabajan con el protocolo OAI y recopilan metadatos Dublin Core y METS. También son destacables las colecciones digitales de todo tipo de documentos desarrolladas por British Library Labs²⁶, muchas de ellas de valor patrimonial. En este caso, la biblioteca ofrece libremente el acceso a los metadatos en diferentes formatos, según cada colección, entre ellos RDF/XML básico, OWL o SKOS con acceso SPARQL (por ejemplo, en la *British National Bibliography*²⁷).

1.4. Herramientas y recursos de interés para explotar las posibilidades de la web semántica

Existen diferentes iniciativas orientadas a facilitar la exploración de las oportunidades y beneficios de publicar datos de repositorios como datos enlazados.

El ciclo de vida de los datos enlazados comprende un conjunto de fases cíclicas, que incluyen la creación de datos, el enlace (reconciliación) con datos RDF externos, el modelado mediante vocabularios/ontologías, y la publicación para facilitar su consumo. Diferentes metodologías han sido

propuestas para concretar y sistematizar dichas fases. En paralelo a la definición del ciclo de vida, se ha desarrollado un sector de herramientas informáticas para la gestión y automatización de fases aisladas, o de varias de ellas de forma conjunta. Una fuente interesante para la identificación y selección de estos tipos de herramientas es Semantic Web Development Tools²⁸ un directorio anotado que es mantenido por la comunidad de usuarios del wiki sobre web semántica del W3C.

Open Refine²⁹ es una herramienta que permite el tránsito por diversas fases del ciclo de vida de los datos enlazados, y que, si bien no ha sido desarrollada para la gestión específica de datos de bibliotecas, archivos y museos, ha tenido un importante impacto en estos ámbitos. A priori, Open Refine facilita la limpieza de grandes cantidades de datos en formatos estructurados, y su enriquecimiento mediante consultas a bases de datos externas. Mediante la incorporación de extensiones como RDF Refine, es posible enlazar los datos originales con datos RDF de conjuntos externos, y transformarlos a RDF mediante el mapeo con vocabularios/ontologías generadas "ad hoc", o importadas. A partir de aquí, es posible exportar los datos transformados a herramientas de explotación e integración de datos RDF.

En el ámbito específico de las bibliotecas y los museos, cabe destacar ALIADA³⁰, que estará disponible como prototipo final en octubre de 2015. Es una aplicación que facilita el tránsito por todas las fases del ciclo de vida de los datos enlazados, desde la transformación de datos de origen en formatos MARC y LIDO, a datos RDF a través de ontologías específicas de esos dos ámbitos (FRBRoo para datos de biblioteca y CIDOC-CRM para datos de museos), hasta su publicación en diversos tipos de plataformas, como, por ejemplo, CKAN. En el momento de escribir este artículo, es posible la descarga e instalación de la versión ALIADA v. 1.0.

Además de estas herramientas, es importante tener en cuenta casos de uso representativos de diferentes recetas de publicación de datos enlazados a partir de colecciones de bibliotecas, archivos y museos. Ya el W3C Library Linked Data Incubator Group realizó una recopilación de estos casos de uso que se incluye en su informe final³¹. Más recientemente, Mitchell (2013) explora tres iniciativas de publicación de datos enlazados orientadas a las comunidades del patrimonio cultural: *Europeana*³², *The Digital Public Library of America*³³, y *BIBFRAME*³⁴, una iniciativa liderada por la Library of Congress. El valor de estos casos es que el análisis de sus similitudes y diferencias permiten inferir buenas prácticas para la aplicación en otros proyectos del sector.

2. OBJETIVOS

Dadas la relevancia de las tecnologías de la web semántica, su relación con los Datos Enlazados y su importancia para las colecciones digitales de entidades patrimoniales, cabe preguntarse cuál es su grado de utilización actual en bibliotecas, archivos y museos. Este trabajo pretende analizar la aplicación del modelo de datos para metadatos RDF en las principales colecciones digitales españolas con una orientación patrimonial y, con ello, ofrecer una visión panorámica del grado de implementación de esta tecnología en las colecciones digitales españolas y, por consiguiente, de su mayor o menor adaptación a la web semántica.

3. METODOLOGÍA

Para evaluar el nivel de aplicación de las tecnologías de la web semántica se ha examinado un conjunto de repositorios con el objetivo de determinar en qué medida expresan sus registros en RDF, ofrecen un punto de consulta SPARQL consultable por agentes externos, y usan referencias como valor de las propiedades, ya sea mediante un URI local o uno externo. En primer lugar, se ha identificado qué programa se emplea en cada repositorio para gestionar la colección digital, ya que el software puede condicionar el modelo de datos. A continuación, se ha averiguado si los registros se expresan en RDF y, en caso afirmativo, si el repositorio ofrece alguna forma de consulta. Por último, se han analizado los resultados de las consultas para describir los modelos mediante los cuales se articulan los datos.

Para la selección de colecciones digitales se han seguido los mismos criterios de selección que en Sulé y otros, 2011; es decir, se han analizado las colecciones recogidas en el directorio de colecciones digitales de *Hispana*³⁵ de carácter exclusivamente patrimonial y que participan en la recolección de datos OAI. Aplicando estos criterios de selección, el 4 de abril de 2014 identificamos un total de 51 colecciones, que se enumeran en un apéndice al final del estudio.

La obtención de los datos de las colecciones digitales se ha realizado aplicando diferentes métodos de análisis según la información que se quería recopilar. La identificación del programa de gestión del repositorio se ha llevado a cabo, principalmente, por medio del análisis directo del sitio web. Cuando esto no ha sido suficiente, se ha enviado un correo electrónico a su responsable solicitando dicha información.

La confirmación, o no, de que los datos contenidos en las colecciones se expresan también según el modelo de datos para metadatos RDF y la

identificación de su forma de consulta (OPAC, API o SPARQL) se ha realizado también por medio de la combinación de la observación directa de los repositorios y del envío de correos electrónicos a los responsables de las colecciones. También se han consultado catálogos de datos abiertos de las administraciones públicas –como son *datos.gov.es*³⁶ y los catálogos específicos de los gobiernos autonómicos y municipales– por si contuvieran los datos RDF de las colecciones.

Por último, los detalles sobre el modelo de datos, las clases y las propiedades utilizados por las colecciones que expresan sus datos en RDF se han obtenido produciendo, cuando así ha sido posible, el fichero RDF/XML correspondiente, generando su DTD con XMLSpy XML Editor³⁷ y analizando los datos resultantes.

A continuación, se presentan los resultados del estudio y su valoración, comentando los programas utilizados en la gestión, si expresan los datos en RDF y cómo pueden consultarse. Se muestran los modelos de datos utilizados, concretamente EDM y OntoWeb, de los cuales se describen las clases y propiedades, para pasar a una evaluación global de la aplicación de estos modelos.

4. RESULTADOS

4.1. Programa de gestión de las colecciones

Los programas empleados para gestionar las 51 colecciones digitales seleccionadas son, ordenados según el número de colecciones que los usan:

- DIGIBIB: 17
- DSpace: 14
- CONTENTdm: 4
- Pandora: 4
- Fedora: 2
- Invenio: 1
- Nubarchiva³⁸: 1
- Microsoft Access: 1
- Desarrollo propio: 5
- Desconocido³⁹: 2

Como se ve, hay un claro predominio de DIGIBIB⁴⁰ (33%) y de DSpace⁴¹ (27%). El primero es un programa propietario de la empresa DIGIBIS orientado a la gestión de colecciones, físicas o digitales, que cuenta con una importante presencia en centros españoles⁴². Basado en MARC21, permite la conversión automática y transparente para el usuario de re-

gistros MARC en Europeana Data Model⁴³ (EDM) para su recolección por *Hispana* y *Europeana*. Por el contrario, DIGIBIB no dispone de ningún punto SPARQL.

Por su parte, DSpace es un programa de código abierto orientado a la gestión de colecciones digitales con una gran implementación internacional en repositorios institucionales. Basado en Dublin Core, DSpace es compatible también con OAI-PMH y desde la versión 5.0 permite convertir y almacenar los datos en RDF, así como publicarlos por medio de un punto de consulta SPARQL o serializados en RDF/XML, Turtle o N-Triples⁴⁴. En relación con Europeana Data Model, vale la pena mencionar el excelente trabajo que está llevando a cabo la Subdirección General de Coordinación Bibliotecaria del Ministerio de Educación, Cultura y Deporte que ha desarrollado el plugin *DSpace2Europeana* que permite a las instalaciones de DSpace interoperar con *Europeana* "a través de añadir la capacidad de exportación de registros de DSpace en los formatos ESE (Europeana Semantic Elements) y EDM (Europeana Data Model)"⁴⁵ (Ibai Sistema, 2012).

4.2. Registros expresados en RDF

En relación con la expresión de los datos en RDF, los resultados obtenidos son:

- Expresan sus registros en RDF: 26
- No expresan sus registros en RDF: 16
- Sin respuesta: 9

De las 26 colecciones que expresan sus registros en RDF, la distribución según el programa de gestión es la siguiente:

- DIGIBIB: 17
- DSpace: 7
- Fedora: 2

Es decir, todas las colecciones gestionadas con DIGIBIB permiten expresar sus registros bibliográficos en RDF, mientras que de las que son gestionadas con DSpace únicamente hemos tenido constancia de siete que lo hacen (de las siete restantes, tres han confirmado que no, y de otras cuatro no hemos obtenido respuesta). Las dos colecciones gestionadas con Fedora⁴⁶ también permiten expresar los registros bibliográficos en RDF.

En relación a los catálogos de datos abiertos de las administraciones públicas (*datos.gov.es* y catálogos específicos de los gobiernos autonómicos y municipales), hemos constatado que, en general, estos depósitos no referencian las colecciones digitales analizadas y que cuando lo hacen no dan cuenta de sus datos RDF.

4.3. Punto de consulta

Respecto a la forma de consulta de los datos expresados en RDF, los resultados obtenidos son:

- SPARQL endpoint: 1
- API (GET method): 21
- OPAC: 3
- Sin respuesta: 1

En este caso, la única colección donde hemos podido constatar que dispone de un punto de consulta SPARQL es el *Fondo histórico digitalizado de la UNED*⁴⁷ gestionado con Fedora, si bien por motivos de seguridad su consulta requiere de nombre de usuario y contraseña, y no hemos podido analizar sus datos.

En cuanto a las 21 colecciones que facilitan la consulta de todos sus datos expresados en RDF a través de una API, el resultado es un fichero XML obtenido mediante una petición "verb=ListRecords" según los parámetros del protocolo OAI-PMH⁴⁸. De todas ellas, catorce corresponden a colecciones gestionadas con DIGIBIB (es decir, no todas) y siete con DSpace.

Por último, hay que mencionar que tres de las colecciones que expresan sus datos en RDF únicamente permiten consultarlos desde el mismo OPAC y de forma individual, es decir, registro por registro. Se trata de tres colecciones gestionadas con DIGIBIB que no disponen de las funcionalidades de publicación en Linked Open Data y en EDM.

Así pues, el análisis de los modelos de datos, las clases y las propiedades de los registros expresados en RDF de las colecciones digitales españolas se ha tenido que limitar a las 21 que lo hacen a través de una API con el método GET. En el resto de casos, como se ha comentado, no hemos obtenido respuesta de la institución, no hemos podido acceder a los datos (SPARQL con claves de acceso) o sólo se pueden consultar los registros individualmente uno por uno (lo que imposibilita garantizar unos resultados exhaustivos para toda la colección).

4.4. Modelo de datos, clases y propiedades

En relación con el modelo en que son expresados los datos, el análisis de los ficheros obtenidos con las consultas GET, realizadas en los meses de mayo y junio de 2015, muestra que la mayoría de las colecciones, 17 de 21, siguen el Europeana Data Model (EDM): 14 de ellas gestionadas con DIGIBIB y tres con DSpace. Las cuatro restantes, gestionadas con DSpace, siguen el extinto modelo *OntoWeb*⁴⁹. En los dos casos, se utiliza XML para la serialización de los datos.

Puesto que ambos modelos establecen especificaciones de clases y de propiedades que pueden ser empleadas de forma diferente, pasamos a analizarlos de forma separada.

4.4.1. Colecciones digitales que siguen el modelo EDM

Como se ha mencionado, son 17 las colecciones que expresan sus datos siguiendo European Data Model. EDM establece tres clases básicas (*core classes*): edm:ProvidedCHO, edm:WebResource y edm:Aggregation. Para nuestro estudio nos interesa la clase edm:ProvidedCHO, puesto que contiene los metadatos que representan a los objetos originales en sí mismos. Sus propiedades obligatorias son:

- dc:title o dc:description
- dc:language para objetos textuales
- dc:subject o dc:type o dc:coverage o dcterms:spatial
- edm:type

A continuación se muestra en la tabla I en cuántas colecciones aparecen las propiedades de la clase edm:ProvidedCHO que se han identificado en las 17 bases de datos analizadas, así como la naturaleza de sus valores. Para una mejor valoración de los resultados, en la columna Datatype se indica el tipo de dato que autoriza EDM para cada elemento⁵⁰.

La primera evidencia es que hay seis propiedades que aparecen en todas las colecciones (columna "No" igual a 0): dc:creator, dc:description, dc:language, dc:title, dc:type y edm:type. De estos seis elementos, cinco corresponden a las propiedades obligatorias mencionadas anteriormente, mientras que dc:creator no. Esta coincidencia viene en gran medida determinada porque cumplir con las especificaciones de EDM es un requisito para que las colecciones puedan integrarse en *Europeana*.

Si nos fijamos ahora en la naturaleza de los valores de las propiedades, y más concretamente en la de aquellas que EDM especifica que puede ser una referencia, vemos que de las 26 posibles únicamente once propiedades se dan en uno o más repositorios con un URI local, destacando por su frecuencia dc:creator, dc:subject, dcterms:spatial, dc:contributor, dcterms:isPartOf y dc:publisher. En algunos casos (dc:creator, dcterms:spatial y dc:Publisher) la mayoría de los repositorios expresan los valores tanto con un URI local o con un literal, dependiendo ello de si el concepto tiene asociado en el catálogo un registro de autoridad o no.

Mención especial merece el caso de la propiedad dc:language en la *Biblioteca Virtual del Patrimonio Bibliográfico* y *IURIS Digital*, puesto que en algunos registros su valor se da en forma de URI externo (en concreto el URL del código ISO 639 de la web de la Library of Congress⁵¹) a pesar de que EDM recomienda el uso de códigos literales de dos o tres letras.

Además de las clases básicas, EDM incluye también cuatro clases contextuales (*contextual classes*): edm:Agent, edm:Place, edm:TimeSpan y skos:Concept. Estas clases contextuales permiten dar información sobre estas entidades de una manera diferenciada del objeto descrito (edm:ProvidedCHO). Los resultados del uso de estas clases contextuales en las 17 colecciones analizadas se pueden ver en la tabla II.

Esta clase aparece en 15 de los 17 repositorios que siguen el modelo EDM. En estas 15 colecciones, la propiedad skos:prefLabel aparece en todas ellas (columna "No" igual a 0), hecho que concuerda con la recomendación de EDM. Tras ella, las más utilizadas son owl:sameAs (nueve colecciones), skos:altLabel (ocho colecciones) y skos:note (seis colecciones). El resto aparecen en tres o menos colecciones.

Por lo que hace a las propiedades que EDM establece que su valor ha de ser una referencia, vemos que su uso no está muy extendido en las colecciones analizadas: owl:sameAs se da únicamente en nueve de ellas y edm:isRelatedTo en tres. En el caso de edm:isRelatedTo los valores son en todos siempre URI locales (referencias a otros agentes de la misma colección). Por el contrario, en la propiedad owl:sameAs hemos encontrado URI externos que referencian al mismo agente en otras bases de datos. En la tabla III se muestran los catálogos referenciados ordenados de mayor a menor frecuencia.

Como se puede observar en la tabla IV, en los repositorios analizados los enlaces externos más frecuentes de la clase contextual Agent son con *VIAF* y con *DBpedia*.

En este caso, la clase contextual Place aparece en catorce de las 17 colecciones analizadas. El análisis también refleja un uso mucho menor de las propiedades establecidas por EDM. De hecho, excepto skos:prefLabel, que se encuentra en las catorce colecciones, el uso de las otras propiedades se reduce a cuatro o menos colecciones.

En las tres colecciones que hemos encontrado la propiedad owl:sameAs, en dos de ellas se referencia a la *Lista de Encabezamientos de Materia para las Bibliotecas Públicas en SKOS* y en una al *Library of Congress Name Authority File* (tabla V).

Tabla I. Propiedades de la clase edm:ProvidedCHO

	Datatype	No	URI local	URI externo	URI local o literal	URI externo o literal	Literal
dc:creator	Literal o ref	0	3	0	12	0	2
dc:description	Literal o ref	0	0	0	0	0	17
dc:language	Literal	0	0	0	0	2	15
dc:title	Literal	0	0	0	0	0	17
dc:type	Literal o ref	0	0	0	0	0	17
edm:type	Literal	0	0	0	0	0	17
dc:date	Literal o ref	1	0	0	0	0	16
dc:publisher	Literal o ref	1	0	0	10	0	6
dc:subject	Literal o ref	2	11	0	3	0	1
dc:contributor	Literal o ref	3	13	0	0	0	1
dcterms:spatial	Literal o ref	3	2	0	12	0	0
dcterms:tableOfContents	Literal	3	0	0	0	0	14
dcterms:temporal	Literal o ref	3	0	0	2	0	12
dc:rights	Literal o ref	4	0	0	0	0	13
dcterms:alternative	Literal	4	0	0	0	0	13
dcterms:isPartOf	Literal o ref	6	8	0	2	0	1
dcterms:isReferencedBy	Literal o ref	8	0	0	0	0	9
dc:relation	Literal o ref	12	0	0	1	0	4
dcterms:replaces	Literal o ref	13	1	0	0	0	3
dcterms:extent	Literal o ref	14	0	0	0	0	3
dcterms:isReplacedBy	Literal o ref	14	1	0	0	0	2
dc:coverage	Literal o ref	15	0	0	0	0	2
dcterms:created	Literal o ref	15	0	0	0	0	2
dcterms:hasFormat	Literal o ref	15	0	0	0	0	2
dcterms:hasPart	Literal o ref	15	2	0	0	0	0
dc:format	Literal o ref	16	0	0	0	0	1
dc:identifier	Literal	16	0	0	0	0	1
dc:source	Literal o ref	16	0	0	0	0	1
dcterms:issued	Literal o ref	16	0	0	0	0	1
dcterms:medium	Literal o ref	16	0	0	0	0	1
edm:currentLocation	Ref	16	0	0	0	0	1
owl:sameAs	Ref	16	0	0	0	0	1

Tabla II. edm:Agent

	Datatype	No	URI local	URI externo	URI local o externo	Literal
skos:prefLabel	Literal	0	0	0	0	15
owl:sameAs	Ref	6	0	8	1	0
skos:altLabel	Literal	7	0	0	0	8
skos:note	Literal	9	0	0	0	6
edm:isRelatedTo	Ref	12	3	0	0	0
rdaGr2:gender	Literal	12	0	0	0	3
edm:begin	Literal	13	0	0	0	2
edm:end	Literal	13	0	0	0	2
rdaGr2:dateOfEstablishment	Literal	14	0	0	0	1
rdaGr2:dateOfTermination	Literal	14	0	0	0	1

Tabla III. edm:Agent / owl:sameAs

VIAF ¹	10
DBpedia ²	7
ISNI ³	2
datos.bne.es ⁴	2
LCNA ⁵	1
FAST Linked Data ⁶	1
Katalog der Deutschen Nationalbibliothek ⁷	1

¹ VIAF: *The Virtual International Authority File*. <https://viaf.org/>. [Consulta: 21/06/2015].

² DBpedia. <http://wiki.dbpedia.org/>. [Consulta: 21/06/2015].

³ ISNI. <http://www.isni.org/>. [Consulta: 21/06/2015].

⁴ DATOS.BNE.ES. <http://datos.bne.es/>. [Consulta: 21/06/2015].

⁵ Library of Congress Name Authority File. <http://id.loc.gov/authorities/names.html>. [Consulta: 21/06/2015].

⁶ FAST Linked Data: *FAST Authority File*. <http://experimental.worldcat.org/fast/>. [Consulta: 21/06/2015].

⁷ Katalog der Deutschen Nationalbibliothek. <https://portal.dnb.de/opac.htm> [Consulta: 21/06/2015].

Tabla IV. edm:Place

	Datatype	No	URI local	URI local o literal	URI local o externo	URI externo	Literal
skos:prefLabel	Literal	0	0	0	0	0	14
skos:altLabel	Literal	10	0	0	0	0	4
owl:sameAs	Ref	11	0	0	1	2	0
skos:note	Literal	12	0	0	0	0	2
dcterms:hasPart	Ref	12	2	0	0	0	0
dcterms:isPartOf	Ref	12	1	1	0	0	0
wgs84_pos:lat	Floating point	13	0	0	0	0	1
wgs84_pos:long	Floating point	13	0	0	0	0	1

Tabla V. edm:Place / owl:sameAs

Lista de Encabezamientos de materia para las Bibliotecas Públicas ¹	2
LCNA	1

¹ *Lista de Encabezamientos de Materia para las Bibliotecas Públicas en SKOS*. <http://id.sgcb.mcu.es/lem/>. [Consulta: 21/06/2015].

La clase edm:TimeSpan es, con mucha diferencia, el elemento de EDM con menor presencia en las colecciones analizadas. Únicamente se encuentra en dos de ellas. Además, en estos dos casos la única propiedad que se incluye es skos:prefLabel.

La última clase analizada es skos:Concept que se encuentra en 15 de las 17 colecciones analizadas. Como se puede ver en la tabla VII, sólo hay una

propiedad (skos:prefLabel) que aparece en todas ellas, mientras que skos:broader lo hace en catorce.

skos:closeMatch, la propiedad que permite dar el URI de un concepto similar de otra base de datos, se encuentra en siete repositorios: en cinco de ellos se referencia a la *Lista de Encabezamientos de materia para las Bibliotecas Públicas*, y en dos a la *Library of Congress Subject Headings* y a la *DBpedia*.

Tabla VI. edm:TimeSpan

	Datatype	No	Literal
skos:prefLabel	Literal	15	2

Tabla VII. skos:Concept

	Datatype	No	URI local	URI local o externo	URI externo	Literal
skos:prefLabel	Literal	0	0	0	0	15
skos:broader	URI	1	14	0	0	0
skos:closeMatch	URI	8	0	1	6	0
skos:altLabel	Literal	10	0	0	0	5
skos:note	Literal	12	0	0	0	3
skos:related	URI	12	3	0	0	0
skos:narrower	URI	14	1	0	0	0
skos:notation	URI	14	0	0	0	1

Tabla VIII. skos:Concept / skos:closeMatch

Lista de Encabezamientos de materia para las Bibliotecas Públicas	5
LCSH¹	2
DBpedia	2

¹ *Library of Congress Subject Headings*. <http://id.loc.gov/authorities/subjects.html>. [Consulta: 21/06/2015].

4.4.2. Colecciones digitales que siguen el modelo OntoWeb

Como se ha comentado en el apartado 4.4, hay cuatro colecciones (todas ellas gestionadas con DSpace) que siguen el modelo de datos OntoWeb. OntoWeb Ontology fue diseñada en el marco del proyecto "OntoWeb: ontology-based information exchange for knowledge management and electronic commerce", financiado por la Unión Europea a través del programa IST (*Tecnologías de la Sociedad de la Información*), entre junio de 2001 y mayo de 2004. La ontología constituía la columna vertebral del portal del mismo nombre, cuyo propósito era integrar los ámbitos de la investigación y la industria para mejorar el intercambio y los servicios de información mediante el uso de ontologías. Al mismo tiempo, OntoWeb también pretendía fortalecer la influencia europea en los procesos de normalización en diferentes áreas de la web semántica.

En su forma original, algunos de los componentes de OntoWeb son utilizados por DSpace en el diseño de una tabla de equivalencias para facilitar la inclusión de datos formalizados en RDF en la exposición de metadatos estructurados a través de OAI-PMH. En este mecanismo se utiliza la clase `ow:Publication` para enmarcar el registro de metadatos descriptivos de los objetos, que se expresan mediante propiedades de Dublin Core.

En ninguna de las cuatro colecciones analizadas en este apartado hemos encontrado clases de contexto (`edm:Agent`, `edm:Place`, `edm:TimeSpan` o `skos:Concept`), por lo que en la tabla IX se muestran únicamente las propiedades de la clase

`ow:Publication` ordenadas, de mayor a menor, por el número de repositorios donde aparecen.

Tabla IX. ow:Publication

	No	Literal
dc:date	0	4
dc:description	0	4
dc:rights	0	4
dc:type	0	4
dc:title	0	4
dc:identifier	0	4
dc:creator	1	3
dc:language	1	3
dc:subject	1	3
dc:publisher	2	2
dc:source	2	2
dc:contributor	3	1
dc:relation	3	1
dc:format	3	1

Como se puede apreciar hay seis propiedades que aparecen en las cuatro colecciones: `dc:date`, `dc:description`, `dc:rights`, `dc:type`, `dc:title` y `dc:identifier`. El resto lo hace con menor frecuencia.

Vale la pena destacar que, en este caso, el valor de todas las propiedades en todas las colecciones es literal, es decir, en ningún caso hemos encontrado un URI (local o externo).

5. CONCLUSIONES

La principal conclusión que se extrae de estos resultados es que la aplicación del modelo de datos RDF en las colecciones digitales analizadas es muy desigual y, en general, poco desarrollada. Únicamente 26 de las 42 colecciones analizadas expresan sus datos en RDF, y de estas 26 la mayoría (17) están gestionadas con DIGIBIB, lo que nos da una idea de hasta qué punto la publicación de datos en RDF puede estar condicionada al uso de aplicaciones que incorporen, ya de serie, funcionalidades para convertir de manera automática sus registros en tripletes RDF.

Un segundo resultado que demuestra el poco desarrollo de la aplicación de tecnologías de la web semántica es la casi inexistencia de puntos de consulta SPARQL con los que permitir que agentes externos puedan consultar y recuperar de manera automática los datos RDF de las colecciones. Únicamente nos consta que una colección, el *Fondo histórico digitalizado de la UNED*, ofrece dicho punto de consulta SPARQL, si bien por motivos de seguridad su consulta requiere de nombre de usuario y contraseña. Del resto de colecciones, en tres casos sólo se pueden obtener los datos RDF consultando directamente el OPAC (registro por registro), mientras que en 21 se puede obtener un fichero RDF/XML con todos los registros expresados en RDF a través de una consulta GET.

Un tercer criterio de evaluación es el uso de referencias como valor de las propiedades, ya sea en forma de URI local o externo. En este caso hay una primera y clara diferencia según el modelo de datos empleado, ya sea éste EDM u OntoWeb. EDM establece para ciertas propiedades que su valor ha de ser o puede ser una referencia, mientras que OntoWeb no establece ningún criterio en este sentido. Esta diferencia de modelos de datos determi-

na que en las cuatro colecciones OntoWeb ninguna propiedad contenga en ningún caso una referencia como valor (siempre es un literal), mientras que en las 17 de EDM encontramos en diversas propiedades valores en forma de URI.

Por lo que concierne al enriquecimiento de los registros con enlaces externos a otras bases de datos, la valoración tampoco es muy positiva. De las 17 colecciones EDM, únicamente nueve dan un URI externo como valor de la propiedad owl:sameAs de la clase edm:Agent, tres en la clase edm:Place y ninguna en la clase edm:TimeSpan. En la clase skos:Concept sólo siete colecciones dan un URI externo como valor de la propiedad skos:closeMatch.

Por todo ello, valoramos que las colecciones digitales españolas deberían hacer un doble esfuerzo enriqueciendo sus datos RDF con enlaces externos y creando puntos de consulta SPARQL para, de esta forma, reforzar sus actuales objetivos de difusión y accesibilidad de las colecciones y, a su vez, desarrollar nuevas metas en el marco de los Datos Abiertos Enlazados.

También sería conveniente determinar los niveles de agregación (local, nacional, europeo) de los datos RDF generados por cada colección. A priori, sería posible considerar únicamente los niveles de la propia colección o del conjunto de las colecciones de una institución, pero, desde nuestro punto de vista, también sería conveniente considerar niveles superiores, como los agregadores nacionales (*Hispana*) o supranacionales (*Europeana*). De esta forma, aumentaría la proyección y difusión de la información contenida en las colecciones. Todo ello, no lo olvidemos, en un momento en que, como ya se ha comentado, los datos y metadatos de bibliotecas, archivos y museos se incorporarán al ámbito de los datos abiertos de la administración del estado.

6. NOTAS

- [1] La versión actualmente vigente es RDF 1.1, cuya especificación está formada por un conjunto de recomendaciones y otros documentos no normativos disponibles en: *RDF current status*. http://w3.org/standards/techs/rdf#w3c_all [Consulta: 21/06/2015].
- [2] Todos los documentos normativos de estas tecnologías y otros documentos informativos para el seguimiento de su evolución pueden consultarse en: *W3C data activity: building the Web of Data*. <http://w3.org/2013/data> [Consulta: 21/06/2015].
- [3] En la versión RDF 1.1 se incorpora otro modelo de identificador, *Internationalized Resource Identifier* o IRI, definido conjuntamente por el World Wide Web Consortium e Internet Engineering Task Force. Resulta de una generalización de URI que permite un margen más amplio de caracteres *Unicode*. Se prevé que progresivamente sustituya a los URI.
- [4] En rigor, el sujeto y el objeto también admiten nodos en blanco (*blank nodes*).
- [5] Las especificaciones de estos lenguajes y otros documentos informativos pueden ser consultados en: *Vocabularies*. <http://w3.org/standards/semanticweb/ontology> [Consulta: 21/06/2015].
- [6] Esta sintaxis de serialización ha sido definida por la recomendación *RDF 1.1 XML Syntax*.

- <http://w3.org/TR/2014/REC-rdf-syntax-grammar-20140225> [Consulta: 21/06/2015].
- [7] Las recomendaciones de W3C que definen estos dos formatos de serialización son, respectivamente, *RDF 1.1 Turtle: Terse RDF Triple Language*. <http://w3.org/TR/turtle> [Consulta: 21/06/2015] y *RDF 1.1 TriG: RDF Dataset Language*. <http://w3.org/TR/2014/REC-trig-20140225> [Consulta: 21/06/2015].
- [8] Las recomendaciones de W3C que definen este lenguaje de serialización son, para XHTML, *XHTML+RDFa 1.1: Support for RDFa via XHTML Modularization*. 3rd ed. <http://www.w3.org/TR/xhtml-rdfa> [Consulta: 21/06/2015], y, para HTML4 y HTML5, *HTML+RDFa 1.1: Support for RDFa in HTML4 and HTML5*. 2nd ed. <http://w3.org/TR/html-rdfa> [Consulta: 21/06/2015].
- [9] Por ejemplo, el formato JSON de RDF (RDF/JSON), cuya sintaxis ha sido definida por la recomendación *RDF 1.1 JSON Alternate Serialization (RDF/JSON): W3C Working Group Note 07 November 2013*. <http://w3.org/TR/rdf-json> [Consulta: 21/06/2015]. No debe confundirse con JSON-LD, que se considera de uso preferente a RDF/JSON.
- [10] La especificación de W3C que define el mecanismo de Microdatos en HTML es HTML Microdata: W3C Working Group Note 29 October 2013. <http://w3.org/TR/microdata> [Consulta: 21/06/2015].
- [11] Véase, por ejemplo, *Microdata to RDF: Transformation from HTML+Microdata to RDF*. 2nd ed. <http://w3.org/TR/microdata-rdf> [Consulta: 21/06/2015]. Incluso se han desarrollado iniciativas para dar apoyo al despliegue y uso del vocabulario Schema.org, con una orientación específica a los Datos Enlazados.
- [12] La especificación de W3C que lo define es *JSON-LD 1.0: A JSON-based Serialization for Linked Data*. <http://w3.org/TR/json-ld> [Consulta: 21/06/2015].
- [13] En concreto, la propiedad de JSON-LD que facilita esta operativa es @context. Más información y ejemplos pueden localizarse en Mitchell (2013, p. 24).
- [14] Este tipo de tripleta es denominada "generalized RDF triple".
- [15] Todos los documentos normativos de estas tecnologías y otros documentos informativos para el seguimiento de su evolución pueden consultarse en: *Linked Data*. <http://w3.org/standards/semanticweb/data> [Consulta: 21/06/2015].
- [16] *Open Bibliographic Data*. <http://opendefinition.org/bibliographic> [Consulta: 21/06/2015].
- [17] Unión Europea. Directiva 2003/98/CE del Parlamento Europeo y del Consejo, de 17 de noviembre de 2003, relativa a la reutilización de la información del sector público. *DOUE-L*, nº 345, de 31.12.2003, p. 90-96.
- [18] España. Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. *BOE*, nº 276, 17 noviembre 2007, p. 47160-47165.
- [19] Unión Europea. Directiva 2013/37/UE del Parlamento Europeo y del Consejo de 26 de junio de 2013 por la que se modifica la Directiva 2003/98/CE relativa a la reutilización de la información del sector público. *DOUE-L*, nº 175, de 27 de junio de 2013, p. 1-8.
- [20] España. Congreso de los Diputados. "Proyecto de Ley por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público". *BOCG. Congreso de los Diputados. Serie A: Proyectos de Ley*, nº 150-1, 30 de abril de 2015, p. 1-11.
- [21] La versión actualmente vigente de SPARQL es 1.1. La especificación de este lenguaje y otros documentos informativos, pueden ser consultados en: *SPARQL corrent status*. http://w3.org/standards/techs/sparql#w3c_all [Consulta: 21/06/2015].
- [22] Es el resultado de la confluencia de Datos Abiertos (interoperabilidad legal) y Datos Enlazados (interoperabilidad técnica).
- [23] *W3C Library Linked Data Incubator Group*. W3C. <http://w3.org/2005/Incubator/ld> [Consulta: 21/06/2015].
- [24] *Kollektionen/Sammlungen*. Zentrales Verzeichnis Digitalisierter Drucke. <https://zvdd.de/dms/browsen> [Consulta: 21/06/2015].
- [25] *Digitale Bibliothek*. Münchener Digitalisierungszentrum. http://www.digitale-sammlungen.de/index.html?c=digitale_sammlungen&l=de [Consulta: 21/06/2015].
- [26] *Digital collections*. British Library Labs. <http://labs.bl.uk/Digital+Collections> [Consulta: 21/06/2015].
- [27] *Free data services*. British Library. <http://www.bl.uk/bibliographic/datafree.html> [Consulta: 21/06/2015].
- [28] *Semantic Web Development Tools*. <http://w3.org/2001/sw/wiki/Tools> [Consulta: 21/06/2015].
- [29] *Open Refine: a free, open source, powerful tool for working with messy data*. Open Refine. <http://openrefine.org/> [Consulta: 21/06/2015].
- [30] *ALIADA: Automatic publication under Linked Data Paradigm of library DATA*. ALIADA. <http://www.aliada-project.eu> [Consulta: 21/06/2015].

- [31] Library Linked Data Incubator Group (2011). *Use Cases*. W3C. <http://w3.org/2005/Incubator/1ld/wiki/UseCases> [Consulta: 21/06/2015]. La información recoge muestras hasta el 2011.
- [32] *Europeana*. <http://europeana.eu> [Consulta: 21/06/2015].
- [33] *The Digital Public Library of America*. Library of Congress. <http://dp.la/> [Consulta: 21/06/2015].
- [34] *BIBFRAME: Bibliographic Framework Initiative*. Library of Congress. <http://loc.gov/bibframe/> [Consulta: 21/06/2015].
- [35] *Hispana: directorio de colecciones digitales*. Ministerio de Educación, Cultura y Deporte. <http://hispana.mcu.es/es/comunidades/directorio.cmd> [Consulta: 21/06/2015].
- [36] *datos.gob.es*. Ministerio de Hacienda y Administraciones Públicas, Ministerio de Industria, Energía y Turismo. <http://datos.gob.es/> [Consulta: 21/06/2015].
- [37] *XMLSpy XML Editor*. Altova. <http://altova.com/xmlspy.html> [Consulta: 21/06/2015].
- [38] Una extensión de Alfresco.
- [39] No se pudo averiguar por la consulta directa del repositorio ni la organización respondió al correo que enviamos solicitando esta información.
- [40] *DIGIBIB para bibliotecas*. DIGIBÍS. <http://www.digibis.com/es/software/digibib.html> [Consulta: 21/06/2015].
- [41] *DSpace*. DSpace. <http://dspace.org> [Consulta: 21/06/2015].
- [42] *Bibliotecas virtuales producidas por DIGIBÍS*. DIGIBÍS. Última actualización: 8 de Enero de 2015. <http://digibis.com/bibliotecas-virtuales/producidas-por-digibis.html> [Consulta: 21/06/2015].
- [43] *Definition of the Europeana Data Model v5.2.5*. 22/05/2014. http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Mapping_Guidelines_v2.2.pdf [Consulta: 21/06/2015].
- [44] *Linked (Open) Data*. DuraSpace Wiki. <https://wiki.duraspace.org/display/DSDO-C5x/Linked+%28Open%29+Data> [Consulta: 21/06/2015].
- [45] *Dspace2Europeana*. Subdirección General de Coordinación Bibliotecaria. <http://www.meecd.gob.es/cultura-mecd/areas-cultura/bibliotecas/recursos-profesionales/8.html> [Consulta: 21/06/2015].
- [46] *Fedora* (<https://getfedora.org>) es un programa de código abierto que permite construir repositorios digitales.
- [47] *Fedora Resource Index Query Service*. <http://62.204.194.43:8080/fedora/risearch> [Consulta: 21/06/2015].
- [48] *Open Archives Initiative Protocol for Metadata Harvesting*. <http://openarchives.org/pmh> [Consulta: 21/06/2015].
- [49] Kim, H. H. (2005). ONTOWEB: implementing an ontology-based web retrieval system. *Journal of the American Society for Information Science and Technology*, 56.11, 1167-1176.
- [50] *Europeana Data Model: mapping guidelines v2.2*. 2014, p. 14-21. http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Mapping_Guidelines_v2.2.pdf [Consulta: 21/06/2015].
- [51] *ISO 639-1: Codes for the Representation of Names of Languages - Part 1: Two-letter codes for languages*. Library of Congress. <http://id.loc.gov/vocabulary/iso639-1.html> y *ISO 639-2: Codes for the Representation of Names of Languages - Part 2: Alpha-3 Code for the Names of Languages*. Library of Congress. <http://id.loc.gov/vocabulary/iso639-2.html> [Consulta: 21/06/2015].

7. REFERENCIAS

- Berners-Lee, T. (2006). *Linked Data*. W3C. <http://w3.org/DesignIssues/LinkedData.html> [Consulta: 21/06/2015].
- Heath, T.; Bizer, C. (2011). *Linked Data: evolving the Web into a global data space* (1st ed.) San Rafael, Ca.; Morgan & Claypool. 136 p. <http://linkeddata-book.com/editions/1.0/> [Consulta: 21/06/2015].
- Ibai Sistemas (2012). *Informe de adaptación de DSpace a Europeana. Fase Danubio: Europeana Data Model (EDM). Parte I: panorama general*. Madrid: Ministerio de Educación, Cultura y Deporte. Subdirección General de Coordinación Bibliotecaria. 91 p. http://travesia.mcu.es/portaln/jspui/bitstream/10421/6301/5/Informe_adaptacion_DSpace_EDM_I.pdf [Consulta: 21/06/2015].
- Ibai Sistemas (2012). *Informe de adaptación de DSpace a Europeana. Fase Danubio: Europeana Data Model (EDM). Parte II: propuesta de solución técnica*. Madrid: Ministerio de Educación, Cultura y Deporte. Subdirección General de Coordinación Bibliotecaria. 32 p. http://travesia.mcu.es/portaln/jspui/bitstream/10421/6301/6/Informe_adaptacion_DSpace_EDM_II.pdf [Consulta: 21/06/2015].

- Kim, H. H. (2005). ONTOWEB: implementing an ontology-based web retrieval system. *Journal of the American Society for Information Science and Technology*, 56 (11), 1167-1176. <http://dx.doi.org/10.1002/asi.20220>
- Library Linked Data Incubator Group (2011). *Library Linked Data Incubator Group Final Report*. W3C. <http://w3.org/2005/Incubator/llid/XGR-llid-20111025> [Consulta: 21/06/2015]
- Mitchell, E. T. (2013). Chapter 2: Building Blocks of Linked Open Data in Libraries. *Library Technology Reports*, 49 (5), 11–25.
- Nilsson, M.; Powell, A.; Johnston, P.; Naeve, A. (2008). *Expressing Dublin Core metadata using the Resource Description Framework (RDF)*. DCMI. <http://dublincore.org/documents/dc-rdf> [Consulta: 21/06/2015]
- RDF 1.1 Concepts and Abstract Syntax (2014). W3C. <http://w3.org/TR/rdf11-concepts> [Consulta: 21/06/2015]
- RDF 1.1 Primer (2014). W3C. <http://w3.org/TR/rdf11-primer> [Consulta: 21/06/2015].
- Suárez-Figueroa, M.C.; Gómez-Pérez, A.; López-Cima, A. (2004). *Portal ontology*. Version 1.0. 36 p. (Deliverable D1.6.2). <http://knowledgeweb.semanticweb.org/semanticportal/deliverables/D1.6.2.pdf> [Consulta: 21/06/2015].
- Sulé Duesa, A.; Estivill Rius, A.; Gascón García, J. (2011). Interfaces de consulta en las colecciones digitales patrimoniales españolas. *Anales de documentación*, vol, 14 (2). <http://revistas.um.es/analesdoc/article/view/113931> [Consulta: 21/06/2015].
- Wood, D.; Zaidman, M.; Ruth, L.; Hausenblas, M. (2014). *Linked Data: Structured Data on the Web*. Shelter Island, NY; Manning. 276 p.
- Yu, L. (2011). *A developer's guide to the semantic web*. Heidelberg; Springer. 608 p. <http://dx.doi.org/10.1007/978-3-642-15970-1>

8. ANEXO

Tabla X. Información sobre las colecciones digitales (datos de febrero de 2015)

Colección digital	URL	Entidad responsable
<i>Arias Montano: repositorio Institucional de la Universidad de Huelva</i>	http://rabida.uhu.es/dspace	Universidad de Huelva
<i>ARCA: arxiu de revistes catalanes antigues</i>	http://www.bnc.cat/digital/arca/	Biblioteca de Catalunya
<i>Biblioteca digital de Castilla y León</i>	http://bibliotecadigital.jcyl.es	Junta de Castilla y León
<i>Biblioteca digital de Castilla-La Mancha</i>	http://bidicam.castillalamancha.es	Biblioteca de Castilla-La Mancha
<i>Biblioteca digital de fondo antiguo</i>	https://biblioteca.unizar.es/colecciones/fondo-antiguo	Universidad de Zaragoza
<i>Biblioteca digital de la ciudad de Murcia</i>	http://murcia.es/bibliotecadigitaldemurcia	Archivo Municipal de Murcia
<i>Biblioteca digital de la Comunidad de Madrid</i>	http://bibliotecavirtualmadrid.org	Bibliotecas de la Comunidad de Madrid
<i>Biblioteca digital de la Real Academia de la Historia</i>	http://bibliotecadigital.rah.es/dgbrah	Real Academia de la Historia
<i>Biblioteca digital Floridablanca</i>	http://bibliotecafloridablanca.um.es/bibliotecafloridablanca	Universidad de Murcia. Biblioteca General
<i>Biblioteca digital hispánica</i>	http://www.bne.es/es/Catalogos/BibliotecaDigitalHispanica/Inicio/index.html	Biblioteca Nacional de España
<i>Biblioteca digital Memoria de Madrid</i>	http://memoriademadrid.es	Ayuntamiento de Madrid
<i>Biblioteca histórica de Santa Cruz</i>	http://www.bhsc.uva.es:8080/BHSC/fondos.jsp	Universidad de Valladolid
<i>Biblioteca navarra digital</i>	https://administracionelectronica.navarra.es/binadi/busqueda.aspx	Servicio de Bibliotecas del Gobierno de Navarra
<i>Biblioteca Saavedra Fajardo de pensamiento político hispánico</i>	http://saavedrafajardo.org	Universidad de Murcia
<i>Biblioteca valenciana digital</i>	http://bv2.gva.es	Biblioteca Valenciana
<i>Biblioteca virtual de Andalucía</i>	http://www.bibliotecavirtualdeandalucia.es	Junta de Andalucía. Consejería de Educación, Cultura y Deporte
<i>Biblioteca virtual de Aragón</i>	http://bibliotecavirtual.aragon.es	Gobierno de Aragón. Dirección General de Cultura

Colección digital	URL	Entidad responsable
<i>Biblioteca virtual de derecho aragonés</i>	http://derechoaragones.es	Gobierno de Aragón
<i>Biblioteca virtual de la provincia de Málaga</i>	http://bibliotecavirtual.malaga.es	Diputación de Málaga
<i>Biblioteca virtual de la Real Academia Nacional de Farmacia</i>	http://bibliotecavirtual.ranf.com	Real Academia Nacional de Farmacia
<i>Biblioteca virtual de La Rioja</i>	http://bibliotecavirtual.larioja.org/bvrioja	Gobierno de La Rioja. Biblioteca de La Rioja
<i>Biblioteca virtual de prensa histórica</i>	http://prensahistorica.mcu.es	Ministerio de Educación, Cultura y Deporte
<i>Biblioteca virtual del patrimonio bibliográfico</i>	http://bvpb.mcu.es	Ministerio de Educación, Cultura y Deporte
<i>Biblioteca virtual del Principado de Asturias</i>	http://www.bibliotecavirtual.asturias.es	Principado de Asturias. Consejería de Cultura y Turismo
<i>Bilboko liburutegi digitala</i>	http://bilbao.net/bld	Ayuntamiento de Bilbao. Red de Bibliotecas Municipales de Bilbao
<i>Calaix</i>	http://calaix.gencat.cat	Generalitat de Catalunya. Departament de Cultura i Mitjans de Comunicació
<i>Cartoteca digital</i>	http://cartotecadigital.icc.cat	Institut Cartogràfic de Catalunya
<i>DARA: documentos y archivos de Aragón</i>	http://www.sipca.es/dara	Gobierno de Aragón. Dpto. de Educación Cultura y Deporte
<i>DIGIBUG: repositorio institucional de la Universidad de Granada</i>	http://digibug.ugr.es	Universidad de Granada
<i>Fondo fotográfico Universidad de Navarra</i>	http://coleccionfff.unav.es/bvunav	Universidad de Navarra
<i>Fondo histórico digitalizado de la UNED</i>	http://e-spacio.uned.es/fez/list.php?community_pid=bibliuned:Fondohistoricodigitalizado	Universidad Nacional de Educación a Distancia
<i>Fondos digitalizados de la Universidad de Sevilla</i>	http://fondosdigitales.us.es	Universidad de Sevilla
<i>Fons Sol-Torres</i>	http://soltorres.udl.cat	Universitat de Lleida
<i>Fototeca del IPCE</i>	http://ipce.mcu.es/documentacion/fototeca.html	Instituto del Patrimonio Cultural de España
<i>Galiciana</i>	http://galiciana.bibliotecadegalicia.xunta.es	Consellería de Cultura, Educación e Ordenación Universitaria da Xunta de Galicia
<i>Hemeroteca Digital</i>	http://hemerotecadigital.bne.es	Biblioteca Nacional de España
<i>IURIS Digital</i>	http://bvrajyl.rajyl.es/i18n/estaticos/contenido.cmd?pagina=estaticos/presentacion	Real Academia de Jurisprudencia y Legislación
<i>Memoria digital de Canarias</i>	http://mdc.ulpgc.es	Universidad de Las Palmas de Gran Canaria
<i>Memòria digital de Catalunya</i>	http://mdc1.cbuc.cat	Consorci de Serveis Universitaris de Catalunya
<i>Memoria digital vasca</i>	http://memoriadigitalvasca.es	Fundación Sancho el Sabio
<i>Patrimonio bibliográfico lacunense</i>	http://hermes.bbtk.ull.es/pandora	Universidad de La Laguna
<i>Proyecto CARMESÍ: archivos históricos de la Región de Murcia</i>	http://www.regmurcia.com/carmesi	Fundación Integra
<i>Recursos de Investigación de la Alhambra</i>	http://www.alhambra-patronato.es/ria	Patronato de la Alhambra y el Generalife. Archivo-Biblioteca
<i>Red digital de colecciones de museos de España: CER.es</i>	http://ceres.mcu.es	Ministerio de Educación, Cultura y Deporte
<i>REGIRA</i>	http://www.ddgi.cat/web/servei.seam?serveiId=1042&nivellId=202&cid=708&categoriaId=2	Diputació de Girona. Servei de Biblioteques
<i>Repositorio de objetos digitales y aprendizaje</i>	http://roda.gobex.es	Junta de Extremadura. Consejería de Cultura
<i>Repositorio documental de la Universidad de Salamanca</i>	http://gredos.usal.es/jspui	Universidad de Salamanca

Colección digital	URL	Entidad responsable
<i>Repositorio documental de la Universidad de Valladolid</i>	http://uvadoc.uva.es	Universidad de Valladolid
<i>Repositorio institucional Arjona y Cubas</i>	http://repositorio.racordoba.es/jspui	Real Academia de Ciencias, Bellas Letras y Nobles Artes de Córdoba
<i>Repositorio institucional de la Universidad de Málaga</i>	http://riuma.uma.es	Universidad de Málaga
<i>Repositorio institucional de la Universidad de Oviedo</i>	http://digibuo.uniovi.es/dspace	Universidad de Oviedo

Colección digital	Programa	RDF	Modelo	SPARQL	API	OPAC
<i>Arias Montano: repositorio Institucional de la Universidad de Huelva</i>	DSpace	Sin respuesta	--	--	--	--
<i>ARCA: arxiu de revistes catalanes antigues</i>	ContentDM	No	--	--	--	--
<i>Biblioteca digital de Castilla y León</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Biblioteca digital de Castilla-La Mancha</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Biblioteca digital de fondo antiguo</i>	Invenio	No	--	--	--	--
<i>Biblioteca digital de la ciudad de Murcia</i>	DSpace	No	--	--	--	--
<i>Biblioteca digital de la Comunidad de Madrid</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Biblioteca digital de la Real Academia de la Historia</i>	DIGIBIB	Sí	EDM	No	No	Sí
<i>Biblioteca digital Floridablanca</i>	DSpace	Sin respuesta	--	--	--	--
<i>Biblioteca digital hispánica</i>	Pandora	Sin respuesta	--	--	--	--
<i>Biblioteca digital Memoria de Madrid</i>	Desarrollo propio	No	--	--	--	--
<i>Biblioteca histórica de Santa Cruz</i>	Desarrollo propio	Sin respuesta	--	--	--	--
<i>Biblioteca navarra digital</i>	Desarrollo propio	No	--	--	--	--
<i>Biblioteca Saavedra Fajardo de pensamiento político hispánico</i>	Desconocido	Sin respuesta	--	--	--	--
<i>Biblioteca valenciana digital</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Biblioteca virtual de Andalucía</i>	DIGIBIB	Sí	EDM	No	No	Sí
<i>Biblioteca virtual de Aragón</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Biblioteca virtual de derecho aragonés</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Biblioteca virtual de la provincia de Málaga</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Biblioteca virtual de la Real Academia Nacional de Farmacia</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Biblioteca virtual de La Rioja</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Biblioteca virtual de prensa histórica</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Biblioteca virtual del patrimonio bibliográfico</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Biblioteca virtual del Principado de Asturias</i>	DIGIBIB	Sí	EDM	No	Sí	Sí

Colección digital	Programa	RDF	Modelo	SPARQL	API	OPAC
<i>Bilboko liburutegi digitala</i>	DSpace	No	--	--	--	--
<i>Calaix</i>	DSpace	Sí	OntoWeb	No	Sí	No
<i>Cartoteca digital</i>	ContentDM	No	--	--	--	--
<i>DARA: documentos y archivos de Aragón</i>	Nubarchiva	No	--	--	--	--
<i>DIGIBUG: repositorio institucional de la Universidad de Granada</i>	DSpace	Sí	OntoWeb	No	Sí	No
<i>Fondo fotográfico Universidad de Navarra</i>	DIGIBIB	Sí	EDM	No	No	Sí
<i>Fondo histórico digitalizado de la UNED</i>	Fedora	Sí	--	Sí	No	No
<i>Fondos digitalizados de la Universidad de Sevilla</i>	Desarrollo propio	No	--	--	--	--
<i>Fons Sol-Torres</i>	DSpace	Sí	OntoWeb	No	Sí	No
<i>Fototeca del IPCE</i>	MSAccess	No	--	--	--	--
<i>Galiciana</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Hemeroteca Digital</i>	Pandora	No	--	--	--	--
<i>IURIS Digital</i>	DIGIBIB	Sí	EDM	No	Sí	Sí
<i>Memoria digital de Canarias</i>	ContentDM	No	--	--	--	--
<i>Memòria digital de Catalunya</i>	ContentDM	No	--	--	--	--
<i>Memoria digital vasca</i>	DSpace	Sí	EDM	No	Sí	No
<i>Patrimonio bibliográfico lacunense</i>	Desconocido	Sin respuesta	--	--	--	--
<i>Proyecto CARMESÍ: archivos históricos de la Región de Murcia</i>	Pandora	Sin respuesta	--	--	--	--
<i>Recursos de Investigación de la Alhambra</i>	DSpace	Sí	OntoWeb	No	Sí	No
<i>Red digital de colecciones de museos de España: CER.es</i>	Desarrollo propio	No	--	--	--	--
<i>REGIRA</i>	Pandora	No	--	--	--	--
<i>Repositorio de objetos digitales y aprendizaje</i>	Fedora	Sí	--	No	Sin respuesta	No
<i>Repositorio documental de la Universidad de Salamanca</i>	DSpace	Sí	EDM	No	Sí	No
<i>Repositorio documental de la Universidad de Valladolid</i>	DSpace	Sí	EDM	No	Sí	No
<i>Repositorio institucional Arjona y Cubas</i>	DSpace	Sin respuesta	--	--	--	--
<i>Repositorio institucional de la Universidad de Málaga</i>	DSpace	No	--	--	--	--
<i>Repositorio institucional de la Universidad de Oviedo</i>	DSpace	Sin respuesta	--	--	--	--



NOTAS Y EXPERIENCIAS / NOTES AND EXPERIENCES

“Desktop Scientometrics”: una metodología para el uso de datos procedentes de WoS mediante el programa estadístico R

Preiddy Efrain-García*, Carlos García-Zorita**

*Centro de Ciencias Humanas y Sociales (CCHS), Consejo Superior de Investigaciones Científicas (CSIC). Madrid, España.

**Laboratorio de Estudios Métricos de la Información (LEMI), Departamento de Biblioteconomía y Documentación.

Universidad Carlos III de Madrid.

Correo-e: preiddy@gmail.com

Recibido: 11-09-2014; 2ª versión: 05-05-2015; Aceptado: 22-05-2015.

Cómo citar este artículo/Citation: Efrain-García, P. y García-Zorita, C. (2016). “Desktop Scientometrics”: una metodología para el uso de datos procedentes de WoS mediante el programa estadístico R. *Revista Española de Documentación Científica*, 39(1): e122. doi: <http://dx.doi.org/10.3989/redc.2016.1.1223>

Resumen: Se presenta una metodología de trabajo que permite obtener y tratar datos de la *Web of Science* con fines bibliométricos. Se usa como herramienta el software estadístico R para crear funciones que faciliten el trabajo bibliométrico; adicionalmente se genera una base de datos relacional que potencia la gestión de los datos y la generación de indicadores bibliométricos.

Palabras clave: Métodos bibliométricos; bases de datos relacionales; SQL; R; Web of Science; indicadores bibliométricos.

“Desktop Scientometrics”: a methodology using the R statistical software with data from WoS

Abstract: A methodology is presented for obtaining and processing data from the Web of Science (WoS) for bibliometric purposes. The R statistical software is used as a tool for creating functions that facilitate bibliometric work. In addition, a relational database was created that enhances both data management and the generation of bibliometric indicators.

Keywords: Bibliometric methods; relational databases; SQL; R; Web of Science; bibliometric indicators.

Copyright: © 2016 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0.

1. INTRODUCCIÓN

Desde que J.S. Katz y Diana Hicks (1997) propusieran el término *Desktop Scientometrics* se ha ido consolidando un "concepto" o "filosofía" de trabajo en relación con los estudios (biblio-) métricos que se ha caracterizado por el uso de herramientas informáticas personales (hardware y software) cada vez más potentes y por la creciente disponibilidad de bases de datos susceptibles de uso bibliométrico.

En estos años se ha llevado a cabo un proceso paralelo en relación con la aparición y desarrollo de aplicaciones de *software libre*, generalmente disponibles en formato multiplataforma, muy accesibles y suministradas en diversos niveles de implementación que posibilitan su uso por diferentes tipos de usuarios.

En este sentido, en el ámbito de los estudios métricos de información es habitual el uso de lenguajes de programación como *Perl* (García González, 2010) o *Python* (Grauwin y Jensen, 2011), la utilización de gestores de bases de datos como *MySQL* (Anuradha y Urs, 2007; Han y Sun, 2014), *Postgre* (Bowman y otros, 2014) o *SQLite* (Gagolewski, 2011), o los análisis llevados a cabo con aplicaciones como *Pajek* (Zhang, y otros, 2013), *CiteSpace* (Chen y otros, 2012) o *Gephi* (Leydesdorff y Rafols, 2012). Todas estas aplicaciones están disponibles para su uso libre y suponen una nueva forma de entender la investigación en la que este trabajo se inscribe.

Por otra parte, la profesión de bibliotecario y en particular la del bibliotecario universitario, se está adaptando en sus competencias a nuevos escenarios. En el caso de las bibliotecas universitarias se está notando un cierto interés en la forma en la que los bibliotecarios pueden integrarse de forma colaborativa en las instituciones que las acogen. Conceptos como el de *'embedded librarian'* (Dewey, 2004) abogan por una mayor integración y colaboración de los bibliotecarios en las tareas de enseñanza-aprendizaje y en las labores de investigación. Tener *'implicados'* profesionales de la información en estos ámbitos académicos puede reportar beneficios a investigadores (Carlson y Kneale, 2011) y docentes (Hernández-Pérez y otros, 2011) para crear nuevas oportunidades a los profesionales (Kesselman y Watstein, 2009).

Así mismo, el creciente interés de las Universidades en las comparaciones nacionales e internacionales *-rankings-* y en la institucionalización de la evaluación de sus profesores en su faceta de investigadores, como ocurre con las acreditaciones o los sexenios en España o el *Research Assessment Exercise* (RAE) en el Reino Unido, han

conducido a un creciente interés en las bibliotecas universitarias por la bibliometría y los servicios que ésta puede prestar en las actividades de evaluación de la investigación de estas instituciones (Corrall y otros, 2013).

Esta nueva apuesta por las competencias bibliométricas / informáticas de los bibliotecarios se ha resuelto incluso contratando expertos en esta área¹ que en el caso de la Universidad de Leicester se hizo incluyendo un bibliómetra en el servicio bibliotecario, sin apostar por la reconversión de sus bibliotecarios (Corrall y otros, 2013).

Sin embargo, en nuestra opinión, y este es el objetivo del presente trabajo, es factible dotar a los profesionales de las bibliotecas universitarias y de investigación de herramientas sencillas que les permitan adquirir y desarrollar competencias básicas que les den la posibilidad de abordar estudios métricos de información desde cualquier computador personal, mediante una metodología basada en el uso de software libre y sin pérdida de control sobre los datos a analizar.

2. BIBLIOMETRÍA CON R

El trabajo bibliométrico que se lleva a cabo hoy en día requiere el empleo de competencias multidisciplinares. Además de las propias de la metodología bibliométrica y como ya se ha insinuado más arriba, son necesarias competencias estadísticas e informáticas. En nuestra propuesta desarrollamos una metodología de tratamiento y análisis de datos que trata de combinar algunos aspectos de estas dos últimas competencias.

R es un conjunto integrado de librerías y un lenguaje de programación que nos permite realizar cálculos fundamentalmente estadísticos (R Core Team, 2012).

R es un entorno en el que se encuentran implementadas muchas técnicas y métodos estadísticos, que van desde los modelos lineales a las series temporales, pasando por los test clásicos. El conjunto básico de instalación se encuentra compuesto por ocho librerías que son llamadas "estándar" (Arriaza Gómez y otros., 2008; Paradis, 2005). Se puede descargar desde su página web principal: <http://www.r-project.org>. Es software libre y se encuentran versiones para los sistemas operativos más empleados en la actualidad².

La comunidad de colaboradores que utilizan el entorno R ha desarrollado centenares de aplicaciones (*packages*) que cubren casi cualquier necesidad de análisis y tratamiento estadístico de datos. En este sentido, y en el ámbito bibliométrico, Gagolewski (2011) ha desarrollado un paquete (CITAN)

para el análisis bibliométrico de citas a partir de datos bibliográficos obtenidos de *Sciverse-Scopus*.

Nuestra intención ahora no es desarrollar un paquete *ad-hoc* para el análisis bibliométrico, sino mostrar, a través de la presentación de algunas funciones sencillas y eficaces, que es relativamente fácil adquirir las competencias necesarias para llevar a cabo análisis más completos.

El entorno de funcionamiento del programa se basa en que posee diversas estructuras de datos denominadas "objetos", a los cuales se le suelen asignar nombres, contenidos y atributos. A continuación mencionamos brevemente algunos de estos objetos:

- *Vector*: Es la estructura más simple, cuenta con dos argumentos *mode* y *length*. Los vectores pueden ser de tres tipos, a saber: numérico, lógico o de carácter. Un vector solo posee un tipo de dato. (R Core Team, 2012; Paradis, 2005).
- *Factor*: "Un factor incluye no solo los valores correspondientes a una variable categórica, sino también los diferentes niveles posibles de esta variable" (Paradis, 2005).
- *Matriz*: Son vectores que cuentan con un atributo adicional (*dim*), que a su vez es un vector numérico de longitud 2 que define el número de las filas y columnas que deberá tener. Las matrices también son denominadas variables indexadas o *arrays* (R Core Team, 2012).
- *Listas*: es un tipo de objeto que está conformado por vectores de diferentes tipos.
- *Dataframes*: son el símil de las bases de datos o las hojas de cálculo. Pueden estar conformadas por diferentes tipos de datos como las listas y más de una columna como en una matriz.

A todos los objetos antes descritos se les pueden aplicar funciones: desde las sencillas de carácter aritmético hasta complejas funciones para el análisis multivariable. La única preocupación que hay que tener presente siempre es con qué tipo de objeto se está trabajando para así poder determinar qué tipo de función podemos aplicar.

Se entiende como función un conjunto de instrucciones escritas en R que reciben un objeto y lo procesan de acuerdo a esas instrucciones. En realidad, todo R trabaja con funciones, es decir, todas las instrucciones que ejecutamos en R son funciones. Para hacernos una idea de qué son las funciones lo mejor es hacer un símil con las macros que se pueden crear en las hojas de cálculo (Paradis, 2005).

Por ser R un lenguaje de programación, el usuario puede programar sus propias funciones. Para ello hay que tener en cuenta algunos aspectos:

- No es necesario *declarar* las variables dentro del entorno de la función, como se hace en otros lenguajes.
- Se escriben directamente en el entorno de trabajo o en un editor externo de texto ASCII.
- Si son escritas directamente en el entorno de trabajo son almacenadas en memoria.
- Si las funciones son escritas con editores externos solo hay que emplear la instrucción *source* (nombre-función) para cargarla en memoria.

En la metodología que aquí presentamos, todos los pasos necesarios para cargar los datos y obtener algunos indicadores han sido programados como una función. Se ha realizado de esta manera para que el usuario final se concentre en lo que consideramos realmente importante, el análisis de los datos y no en aspectos más técnicos. Sin embargo, todo el código de las funciones que empleamos se encuentran en texto plano y pueden ser modificadas y mejoradas con el único compromiso de comunicarlo a los autores.

Ilustramos con un ejemplo la creación de una función y seguidamente se comentan las tareas que realiza.

```
mifuncion=function(x){datos= read.table(x)
plot(datos$V1)}
```

En la primera línea se declara el nombre que tendrá la función (*mifuncion*) para seguidamente entrar en el entorno *function*. En la segunda línea se indica al programa que lea un archivo de texto cuyo nombre se ha pasado como valor (*x*) de la función, el comando *read.table* lee el fichero y su contenido lo almacena en una variable, en este caso del tipo *dataframe*, denominada *datos*. La última línea realiza la correspondiente gráfica (*plot*) de la primera variable (columna *V1*) del archivo indicado en el argumento.

No se pretende que el usuario sea un programador experimentado, sino que disponga de las nociones básicas del funcionamiento de R y que en la medida que el modelo de trabajo que proponemos le resulte útil y motivador, se anime a modificar y crear sus propias funciones.

3. MATERIAL Y MÉTODO

En el método bibliométrico habitualmente son dos las fases a las que hay que enfrentarse: la preparación de los datos y la obtención de indicadores.

Mostramos a continuación cómo afrontar estas dos fases utilizando un conjunto de funciones programadas en el entorno R. El lector podrá ejecutarlas simplemente y si lo desea revisarlas y adaptarlas a sus necesidades.

3.1. Preparación y Obtención de los Datos

En esta primera fase, han de seguirse los pasos necesarios para la obtención y el posterior tratamiento de los datos bibliográficos susceptibles de uso bibliométrico. Una vez seleccionada la fuente de información y obtenidos los datos objeto del estudio es aconsejable la carga, tratamiento e integración de los datos en un sistema de gestión de bases de datos, preferentemente de carácter relacional. Nuestra elección en este trabajo es la *BD SQLite* debido a su sencillez de manejo (un único fichero por cada base de datos). Gestores de *SQLite* hay muchos disponibles libremente y para todos los sistemas operativos, aunque por su simplicidad de instalación recomendamos el complemento disponible para *Firefox*³. Esta fase debe concluir finalmente con la depuración y homologación de los datos.

i. Selección

En el presente trabajo y para ilustrar la metodología propuesta se utilizarán los índices de citas de la *Web of Science (Thomson-Reuters)*; los resultados se almacenarán en archivos de texto plano con extensión “.txt”. Se ha de seleccionar el formato delimitado por tabuladores. Para evitar, en la medida de lo posible, errores en la importación y por la experiencia de trabajo con plataformas Windows y Mac, sugerimos que los datos sean exportados con el tabulado correspondiente a la plataforma que estemos usando y con la codificación de caracteres UTF-8.

ii. Carga de los datos

Para ilustrar las distintas fases metodológicas, y a modo de ejemplo, en este trabajo se ha trabajado con una sola descarga de 500 registros del *Social Sciences Citation Index (SSCI- WoS)*, pero el usuario puede trabajar con el número de ficheros que desee.

Una vez obtenido el o los archivos con el formato delimitado por tabuladores, iniciamos R y procedemos a cargar las funciones que hemos desarrollado, las cuales se encuentran disponibles en la siguiente dirección web: https://github.com/06122010/r_isi⁴

Procedemos a cargar en el entorno de trabajo en R las funciones. A continuación mostramos la instrucción para cargar todas las funciones (no olvide la ruta de acceso al fichero ni las comillas)⁵.

```
> source("../script/funciones.R")
```

Comentaremos, brevemente, las funciones que han sido cargadas en esta primera versión y los parámetros de cómo han de ser usadas en todo momento.

Todas las funciones tienen el prefijo *isi*, seguido de guión bajo y el objetivo por la que fue creada. La primera función es *isi_cargar*, como su nombre indica será la encargada de leer el o los archivos de texto y añadirlos a un *dataframe*. La forma de usar esta función es la siguiente:

```
WoS <- isi_cargar("../txt/")
```

Lo que le estamos diciendo a la función es que nos cargue en un *dataframe* llamado “WoS” todos los archivos que se encuentran en el directorio “../txt/” (no olvide ni las comillas ni el *slash* final). El usuario indicará la ruta donde tenga almacenado todos los .txt de su descarga. Puede nombrar el *dataframe* como desee, no es necesario que se llame WoS, lo usamos como ejemplo.

El proceso de carga suele ser un proceso que consume muchos recursos de nuestro ordenador y por tanto dependiendo del volumen de datos se puede volver muy lento.

Una vez que la función termina el proceso, tendremos en nuestro espacio de trabajo el *dataframe* con todos los datos cargados; ahora bien, para continuar disponemos de dos posibilidades: seguir trabajando con los datos cargados en la memoria o exportar los datos a una base de datos. Para esta segunda opción es necesario instalar en R ciertas librerías: dentro de la propia función se ha agregado verificación de que la librería se encuentra instalada y de no ser así, ésta la instala. Las librerías necesarias para continuar el proceso son: *plyr*; *DBI*; *RSQLite* y *ggplot2*.

Para exportar los datos a un sistema de bases de datos relacionales emplearemos la función ‘*isi_exportar_sqlite*’. Como su nombre indica, exporta el *dataframe* que le indiquemos a una base de datos SQLite. Esta función trabaja con dos parámetros, el primero ha de ser el nombre del *dataframe* que queremos exportar y el segundo la tabla donde queremos almacenar el contenido. La función se encuentra preparada para exportar a tablas los siguientes datos: el *dataframe* principal almacenado en la tabla que hemos llamado ‘*isi_t_pg*’; los autores en la tabla ‘*isi_t_au*’; las direcciones en la tabla ‘*isi_t_c1*’; las categorías temáticas en la tabla ‘*isi_t_wc*’ y finalmente las referencias bibliográficas en la tabla ‘*isi_t_cr*’. La base de datos la hemos llamado ‘*pgwos.db*’ y se creará en la raíz del disco o en la carpeta del proyecto.

Para exportar el *dataframe* WoS a la tabla principal de la base de datos ejecutamos la siguiente instrucción:

```
isi_exportar_sqlite(WoS,"WoS")
```

Si el usuario opta por continuar sin guardar el *dataframe* principal (**WoS**) en una BD, este podrá comenzar a obtener indicadores de los denominados unidimensionales, tales como frecuencias de datos referidos a la fecha, tipología documental, entre otros. Por ejemplo, una sencilla tabla de distribución de frecuencias de la variable fecha en el *dataframe* principal (**WoS\$PY**), podría obtenerse con las siguientes instrucciones:

```
> q=as.data.frame(table(WoS$PY))
> q=cbind(q,F.Acum=cumsum(q$Freq))
```

Cuyo resultado se presenta de la siguiente forma:

PY	Freq	F.Acum
2004	9	9
2005	16	25
2006	14	39
2007	22	61
2008	25	86
2009	26	112
2010	35	147
2011	31	178
2012	40	218
2013	18	236

La instrucción **table** genera una tabla de frecuencias (**PY**, **Freq**) de la variable elegida que se almacena en el *dataframe* (**q**). La instrucción **cbind** añade columnas a este *dataframe* y lo renombra igualmente (**q**), el argumento **cumsum** calcula la frecuencia acumulada de la variable (**PY**). Se anima al lector a probar con otras variables de ocurrencia simple -un único dato por registro- como las mencionadas más arriba o, por ejemplo, la referida al título de la fuente de publicación. Sin embargo, para realizar otros tipos de estudios, basados en campos de ocurrencias múltiples, es necesario procesar dichos datos separándolos en otros subconjuntos de datos. Este es el caso de los listados de autores, de direcciones, y categorías temáticas, entre otras. A continuación describimos el procedimiento para tal fin.

Crearemos subconjuntos de datos referentes a los autores, las direcciones, las categorías y las referencias bibliográficas (si éstas se han incluido en las descargas de datos originales). Para ello, es necesario emplear la función que hemos denominado

'isi_separar', que admite como argumento alguno de los siguientes: 'C1' para crear el *dataframe* con las direcciones; 'AU' para los autores; 'WC' o 'SC' para las categorías temáticas y 'CR' para las referencias bibliográficas.

A continuación crearemos cada uno de los *dataframes* y luego los almacenamos en sus correspondientes tablas en la BD.

```
#Instrucciones que crean los diferentes Dataframes
autores <- isi_arar(WoS,"AU")
direcciones <- isi_separar(WoS,"C1")
temas <- isi_separar(WoS,"WC")
categorias <- isi_separar(WoS,"SC")
referencias <- isi_separar(WoS,"CR")
#Almacenamos en la BD
isi_exportar_sqlite(autores,"AU")
isi_exportar_sqlite(direcciones,"C1")
isi_exportar_sqlite(temas,"WC")
isi_exportar_sqlite(categorias,"SC")
isi_exportar_sqlite(referencias,"CR")
```

Para dar por concluida esta primera parte, se explica brevemente cómo almacenar los datos en otra fuente externa, para su posterior tratamiento, como puede ser una hoja de cálculo.

```
write.table(WoS, "wos.txt", sep="\t")
```

La instrucción anterior salva los datos en un archivo '.txt' separado por tabulaciones. Aunque también podemos usar:

```
write.csv(WoS, file = "wos.csv")
```

Si queremos salvar los datos en una hoja de Excel es necesario instalar y cargar la librería "xlsx"

```
library(xlsx)
write.xlsx(WoS, "wos.xlsx")
```

Adicionalmente podemos instalar y usar la librería "foreign" que permite importar y exportar datos a SPSS, SAS o Stata.

```
library(foreign)
write.foreign(WoS, "wos.txt", "wos.sps",
package="SPSS")
```

3.2. Obtención de indicadores

En este apartado mostramos algunas de las posibilidades de la metodología que aquí proponemos. Se ha programado, a modo de ejemplo, una función, 'isi_unimetrics_wos', que emplea la potente librería gráfica 'ggplot2' para obtener directamente algunos indicadores unidimensionales gráficos desde el *dataframe* principal (WoS).

La función "isi_unimetrics_wos" acepta dos parámetros para su funcionamiento, el primer parámetro es el nombre del *dataframe* principal y el se-

gundo argumento es el campo objeto del análisis. Se pueden analizar, en esta versión de la función, el campo PY referente a la fecha de publicación de los trabajos; el campo DT que muestra la tipología documental; LA para los idiomas. Se ha añadido un parámetro adicional, ACUM, para tener una tabla de frecuencias por fecha y el crecimiento acumulado. De este modo, la sintaxis `isi_unimetrics_wos(WoS,"PY")` obtiene como resultado un gráfico como el de la figura 1.

Para mostrar la flexibilidad de nuestra propuesta, se ha programado una función más, denominada 'h_index' que calcula el *índice-h* de la distribución en un año determinado a partir de los datos almacenados en el *dataframe* WoS. La sintaxis de ejecución es:

```
h_index(WoS,"PY","2010")
```

```
## =====
## El índice h para el 2010 es: 7
## =====
```

La función admite un intervalo de fechas (i), para lo cual hay que ejecutarla con la siguiente sintaxis:

```
for(i in 2010:2014){h_index(WoS,"PY", i)}
```

Si se quiere calcular el *índice-h* en una determinada disciplina, utilizaremos la función del R-core, 'subset', con la siguiente sintaxis:

```
sub_WoS=subset(WoS, grepl("Physics*",WoS$WC))
```

el nuevo *dataframe* (sub_WoS) podrá usarse con

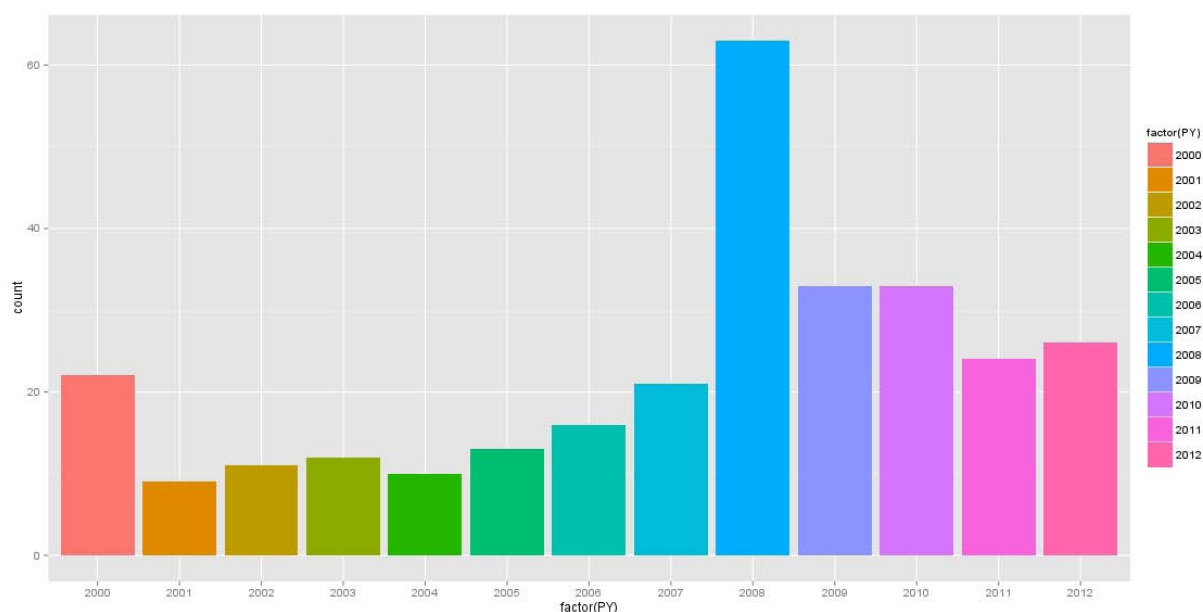
la función "h_index", para calcular el *índice-h* (en un año o en un periodo). El argumento **grepl** del comando **subset**, permite aplicar expresiones regulares para recuperar el área temática correspondiente en el campo WC del *dataframe* principal (WoS).

4. CONCLUSIONES

Sabemos que hay soluciones técnicamente mejores y más elegantes en su planteamiento y ejecución, pero lo que se persigue en este trabajo es presentar una colección de recursos asequibles y con los menores requerimientos informáticos y estadísticos posibles. Recursos, que no solo estén disponibles libremente, condición que nos autoimponemos, sino que sean fácilmente utilizables por quienes necesiten llevar a cabo estudios bibliométricos rigurosos. En particular hemos pensado que será útil a aquellos bibliotecarios que vayan asumiendo las nuevas funciones de apoyo y evaluación de la investigación. Complementariamente la metodología que se expone en este trabajo puede ser de interés para aquellos estudiantes de Información y Documentación que vayan a defender su Trabajo Final de Grado (TFG) en el ámbito de los estudios métricos de información.

Se presentan herramientas autónomas, que funcionan sin apenas control del usuario, pero la filosofía que se persigue es la de motivar a los lectores en la exploración de opciones más avanzadas tratando de convencer que con un mínimo esfuerzo, y desde luego una mayor motivación, pueden ad-

Figura 1.



quirirse las competencias necesarias para realizar análisis bibliométricos y cuantitativos fiables.

En la obtención de datos susceptibles de uso bibliométrico el formato `.csv` se ha convertido en un estándar *de facto*, con muchas alternativas de tratamiento. Las hojas de cálculo lo han sido, y en gran parte lo siguen siendo, pero aunque han mejorado en sus limitaciones globales, como el número de filas o columnas a tratar, aún siguen teniendo grandes limitaciones, como el número de caracteres que se admite en las celdas (255), y si tenemos valores en los campos que sobrepasen ese número no serán tratados correctamente, pensemos en algo cada vez más habitual como los trabajos *hiper-autorados* con centenares e incluso miles de autores y sus correspondientes direcciones institucionales.

El usar una base de datos relacional en este método nos aporta las bondades del lenguaje SQL,

con lo que podemos organizar los datos de una forma más eficiente y con un mayor potencial en su tratamiento que lo que puede resultar del uso de gestores bibliográficos o de las herramientas de análisis incorporadas en las fuentes bibliográficas utilizadas. Hay vías que sólo se insinúan en este trabajo, como es la posibilidad de la consulta directa, sin cliente o gestor de base de datos, desde herramientas globales como es el propio software R.

Por último, deseamos hacer notar que lo que proponemos en este trabajo es una adaptación a los nuevos tiempos de la vieja filosofía del *'desktop scientometrics'*, esta vez basada en aplicaciones multiplataforma y de Software Libre que permita convencer a los usuarios de que con pequeños esfuerzos y mentalidades abiertas al cambio tecnológico es posible disponer en la palma de su mano de herramientas muy potentes para elaborar una amplia variedad de indicadores de interés bibliométrico.

5. NOTAS

- [1] La Universidad de Leicester fue la primera en el Reino Unido en contratar a un especialista en bibliometría (<http://www2.le.ac.uk/news/blog/2010-archive/march-2010/first-uk-academic-bibliometrician-appointed>).
- [2] Por su utilidad recomendamos el entorno de trabajo RStudio™ (<http://www.rstudio.com/>) que es una interfaz integrada libre (no la única) para trabajar con R de una manera más cómoda y amigable.
- [3] Disponible en <https://addons.mozilla.org/es/firefox/addon/sqlite-manager/>.

- [4] Además del *script* con las funciones que se presentan en este trabajo ("funciones.r"), está disponible a modo de guía práctica una versión reducida de este trabajo ("README.md").
- [5] Si usa el *IDE RStudio*, lo más recomendable es iniciar un proyecto nuevo y crear dentro de él una estructura de carpetas. En nuestro caso el fichero está en una carpeta denominada `/script/`, y los ficheros descargados de la WoS en otro denominado `/txt/`. Si se usa esta opción las rutas deberán comenzar por `../`.

6. REFERENCIAS

- Anuradha, K.T.; Urs, S.R. (2007). Bibliometric indicators of Indian research collaboration patterns: A correspondence analysis. *Scientometrics*, vol. 71, no. 2, pp. 179-189. <http://dx.doi.org/10.1007/s11192-007-1657-4>
- Arriaza Gómez, A. J.; Fernández Palacín, F.; López Sánchez, M. A.; Muñoz Márquez, M.; Pérez Plaza, S.; Sánchez Navas, A. (2008). *Estadística básica con R y R-Commander* [en línea]. Cádiz: Universidad de Cádiz, Servicio de Publicaciones. [Consulta: 10 enero 2015]. Disponible en: <http://gsyc.escet.urjc.es/~herraiz/ebrcmdr.pdf>
- Bowman, T.D.; Tsou, A., Ni; C.; Sugimoto, C.R. (2014). Post-interdisciplinary frames of reference: exploring permeability and perceptions of disciplinarity in the social sciences. *Scientometrics*, vol. 101, no. 3, pp. 1695-1714. <http://dx.doi.org/10.1007/s11192-014-1338-z>
- Carlson, J.; Kneale, R. (2011). Embedded librarianship in the research context: Navigating new waters. *College & Research Libraries News*, vol. 72, no. 3, pp. 167-170.
- Chen, H.; Zhao, G.; Xu, N. (2012). The Analysis of Research Hotspots and Fronts of Knowledge Visualization Based on CiteSpace II. En: S.K.S. Cheung, J. Fong, L.-F. Kwok, K. Li y R. Kwan (eds.), *Hybrid Learning* [en línea]. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 57-68. [Consulta: 10 enero 2015]. Disponible en: http://link.springer.com/10.1007/978-3-642-32018-7_6.
- Corrall, S.; Kennan, M.A.; Afzal, W. (2013). Bibliometrics and Research Data Management Services: Emerging Trends in Library Support for Research. *Library Trends*, vol. 61, no. 3, pp. 636-674. <http://dx.doi.org/10.1353/lib.2013.0005>
- Dewey, B.I. (2004). The Embedded Librarian. *Resource Sharing & Information Networks*, vol. 17, no. 1-2, pp. 5-17. http://dx.doi.org/10.1300/J121v17n01_02

- Gagolewski, M. (2011). Bibliometric impact assessment with R and the CITAN package. *Journal of Informetrics*, vol. 5, no. 4, pp. 678-692. <http://dx.doi.org/10.1016/j.joi.2011.06.006>
- García González, P.E. (2010). *Diseño, desarrollo y aplicación de un método para el análisis y tratamiento de la información con fines métricos*. [Tesis doctoral]. Universidad Carlos III de Madrid.
- Grauwin, S.; Jensen, P. (2011). Mapping scientific institutions. *Scientometrics*, vol. 89, no. 3, pp. 943-954. <http://dx.doi.org/10.1007/s11192-011-0482-y>
- Han, G.; Sun, B. (2014). Design and Implementation of Bibliometrics System Based on RIA. En: S. Patnaik, S.; Li, X. (editores), *Proceedings of International Conference on Soft Computing Techniques and Engineering Application* [en línea]. New Delhi: Springer India, pp. 263-272. [Consulta: 10 enero 2015]. Disponible en: http://link.springer.com/10.1007/978-81-322-1695-7_30
- Hernández-Pérez, T.; Pacios, A.R.; Vianello, M.; Ortega, R.; Gorospe, M.R. (2011). La formación en alfabetización en información en las aulas universitarias: el caso de la UC3M. *Scire: representación y organización del conocimiento*, vol. 17, no. 2, pp. 27-37.
- Katz, J.; Hicks, D. (1997). Desktop Scientometrics. *Scientometrics*, vol. 38, no. 1, pp. 141-153. <http://dx.doi.org/10.1007/BF02461128>
- Kesselman, M.A.; Watstein, S.B. (2009). Creating Opportunities: Embedded Librarians. *Journal of Library Administration*, vol. 49, no. 4, pp. 383-400. <http://dx.doi.org/10.1080/01930820902832538>
- Leydesdorff, L.; Rafols, I. (2012). Interactive overlays: A new method for generating global journal maps from Web-of-Science data. *Journal of Informetrics*, vol. 6, no. 2, pp. 318-332. <http://dx.doi.org/10.1016/j.joi.2011.11.003>
- Paradis, E. (2005). *R for Beginners* [en línea]. [Consulta: 10 enero 2015]. Disponible en: http://cran.r-project.org/doc/contrib/rdebuts_es.pdf
- R Core Team. (2012). *R: A Language and Environment for Statistical Computing* [en línea]. Vienna, Austria: R Foundation for Statistical Computing. [Consulta: 10 enero 2015]. Disponible en: <http://www.R-project.org>
- Zhang, L.; Thijs, B.; Glänzel, W. (2013). What does scientometrics share with other «metrics» sciences? *Journal of the American Society for Information Science and Technology*, vol. 64, no. 7, pp. 1515-1518. <http://dx.doi.org/10.1002/asi.22834>



NOTAS Y EXPERIENCIAS / NOTES AND EXPERIENCES

Construction of a taxonomy for medieval Portuguese history: problems and challenges

Filipa Medeiros*, Hermínia Vasconcelos Vilar**, José Antonio Moreiro González***

* Centro Interdisciplinar de História, Culturas e Sociedades da Universidade de Évora / Campo Arqueológico de Mértola, Portugal.

** Departamento de História, Universidade de Évora / Centro Interdisciplinar de História, Culturas e Sociedades da Universidade de Évora.

*** Departamento de Biblioteconomía y Documentación, Universidad Carlos III de Madrid.

Correo-e: fmcm@uevora.pt, hmav@uevora.pt, jamore@bib.uc3m.es

Recibido: 28-11-2014; 2ª versión: 13-02-2015; Aceptado: 08-05-2015.

Cómo citar este artículo/Citation: Medeiros, F.; Vilar, H. V.; Moreiro González, J. A. (2016). Construction of a taxonomy for medieval Portuguese history: problems and challenges. *Revista Española de Documentación Científica*, 39(1): e123. doi: <http://dx.doi.org/10.3989/redc.2016.1.1243>

Abstract: Our main goal was to design and build a taxonomy of medieval Portuguese history using an interdisciplinary approach based on doctoral research. First, the criteria used for the selection of the vocabulary and its formal and semantic normalization were determined. Then species were listed, followed by the characterization of categories and their respective subclasses. As conclusions we highlight the successful application of the selected terms, as well as the fact that the taxonomy's categories are being continuously updated and expanded, both in their global extension and in the depth of their thematic representation. In addition we offer proposals for continuing the ontological development of this taxonomy.

Keywords: Taxonomies; KOS; information representation; medieval Portuguese history; interdisciplinarity.

Sobre la elaboración de una taxonomía de Historia medieval portuguesa: problemas y desafíos

Resumen: Se ha fijado como objetivo principal diseñar y elaborar una taxonomía sobre Historia medieval portuguesa, y hacerlo mediante un abordaje interdisciplinar y como fruto de las investigaciones de tesis. En su proceso se determinaron, primero, los criterios seguidos para la selección y normalización formal y semántica del vocabulario. Luego se listaron las especies y se caracterizaron las categorías y las respectivas subclases. Como conclusión se destaca el hecho de que se han experimentado con éxito métodos antes ya probados en otros estudios para proyectar, desarrollar y mantener taxonomías, con independencia del nivel de especificidad temática, hecho que plantea otro tipo de retos. Sin olvidarse de avanzar propuestas que le den continuidad hacia su desarrollo ontológico.

Palabras clave: Taxonomías; KOS; representación de la información; historia medieval portuguesa; interdisciplinarietàad.

Copyright: © 2016 CSIC. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0 License

1. INTRODUCTION

This note describes a taxonomy aimed at organizing and representing information related to Portuguese Medieval History between the 12th and the 15th centuries, which is currently available online as a structured list. This is an applied work (Medeiros, 2014), adapting methods that have already been tested by experts in order to build the taxonomy.

The construction of a taxonomy within the scope of Medieval History is, as we shall see, a pioneering choice and, as such, there are no parameters available for comparison. It should be understood as an instrument that undergoes a continuous construction and evaluation process, but whose purposes are clear.

The construction of this taxonomy seeks to fill a gap found in the Portuguese information units specialized in Medieval History, many of which are integrated into universities and their research centres. Moreover, the construction of this controlled vocabulary seeks mainly to address the need for the indexation of digital resources, namely the ones contained in specialized databases, which are currently seen as crucial vehicles for disseminating and sharing open-access scientific knowledge. We can mention, as an example, a series of databases that have already been developed by the Institute for Medieval Studies of the Faculdade de Ciências Sociais e Humanas of the Universidade Nova de Lisboa (IEM - FCSH/UNL), available at: <http://iem.fcsh.unl.pt/section.aspx?kind=noticia&id=49>

2. METHODS AND RESOURCES

In order to build and develop this taxonomy we used a top-down, or general-to-specific, method (Conway and Sligar, 2002; Cumming, 2003; Gilchrist, 2003; Jagerman, 2006; Zhonghong et al, 2006; Moreiro González, 2006; Moreiro González, 2011), which is the most commonly used method for the construction of this type of controlled vocabulary. This method comprises the following stages:

- a) Compiling the knowledge;
- b) Reducing the number of synonyms and choosing the preferential terms;
- c) Preparing a preliminary systematic structure;
- d) Developing the first draft of the scheme;
- e) Completing the taxonomy with all its terms;
- f) Evaluating its performance;
- g) Publishing its first version.

Next, we detail the specific development of each of the stages in the construction of the Portuguese Medieval History taxonomy, following the top-down method.

a) Compilation of knowledge: in this first stage we surveyed the available specialized sources of information, focusing mainly on vocabularies, lexicons and dictionaries dealing with the Middle Ages. We did not find any documents of Portuguese origin, so we collected foreign reference documents within the scope of general Medieval History that were likely to enrich the taxonomy, specifically in what regarded the definition of its general categories, both in terms of form and in terms of content. We reached a total of 11 foreign sources of information. In the absence of Portuguese sources of information, we resorted to national general histories and university textbooks of reference within the scope of Portuguese medieval studies (Serrão and Marques, 1987; Marques, 1988; Tavares, 1990; Coelho, 1991; Branco and Costa, 1992; Tavares, 1992; Serrão, 1994; Moreno, 1995; Serrão and Marques, 1996; Mattoso, 1997; Ramos et al, 2009).

b) Reduction of synonyms and choice of the preferential terms: this item corresponds to the second stage in the construction of the taxonomy, in which we set out the issues concerning the formal and semantic control of the vocabulary, according to the following analytical scheme shown in Figure 1.

It should be noted here that since the source language of the taxonomy is Portuguese, we naturally followed the guidelines provided by national standards, namely the 4036 Portuguese Standard (items 6 and 7) and the *Siporbase: sistema de indexação em português* manual (section 4, Terminology) (NP 4036, 1992; Portugal, 1998).

The morphological and syntactic control has not proved particularly complex. The Portuguese language rules were followed, as well as the national standards. The same did not occur with the semantic control of some terms, as is the case of the word "Jantar" and the term "Cantigas de escárnio e de maldizer". The first is an example of multiple meanings (semantic ambiguities via significant pathway) and the second, a case of synonymy (semantic ambiguities via meaning).

As we can see in figure 2, and according to NP 4036 (NP 4036, 1992: 13), for these cases we resorted to the use of qualifiers designated in brackets, which are intended to clarify the conceptual content to which the term refers by establishing its meaning, thus avoiding ambiguities in information retrieval.

Figure 1. Stages of formal and semantic normalization

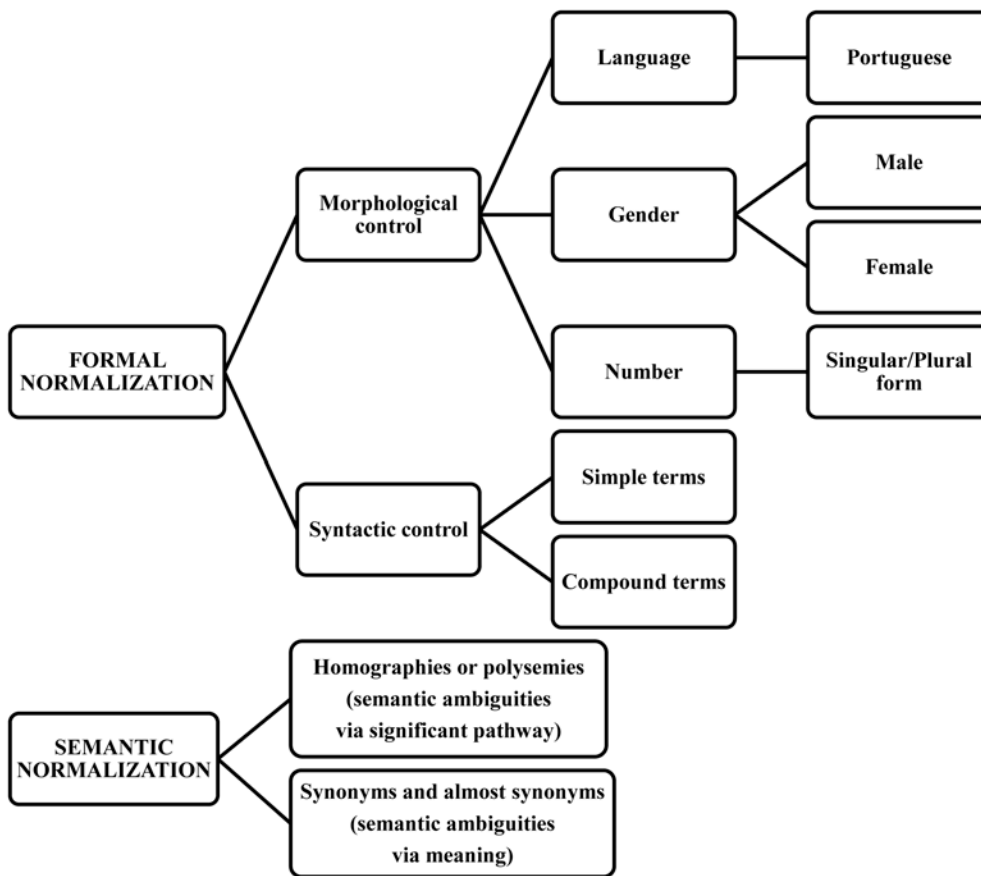


Figure 2. Formal and semantic normalization of the term "Jantar"

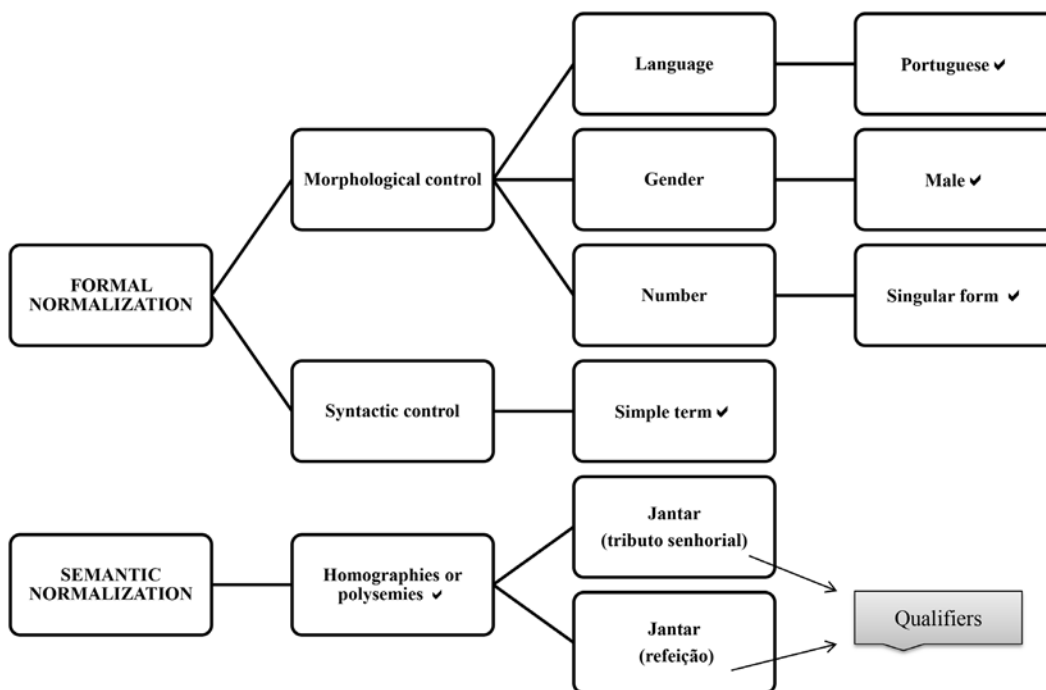


Figure 3 presents a case of synonymy between the terms “Cantigas” and “Canções”. In this situation, our attention was focused on the term generally used by the scientific community (preferred terms), thereby facilitating the compatibility between natural language and the controlled terms present in the taxonomy.

The distribution of the terms across the hierarchical structure was made alphabetically, except in the «Events», «Personalities» and «Reigns» categories, in which the terms were introduced chronologically. This was the only way to organize them diachronically, allowing the user to browse the taxonomic structure according to the natural sequence of events, personalities and reigns, respectively:

Ex. 1128, Batalha de São Mamede

Ex. 1350-1405, Leonor Teles

Ex. 1385-1433, Reinado de D. João I

c) Construction of the taxonomy: After the formal and semantic normalization of the vocabulary, which followed the above-mentioned criteria and procedures, we recorded the compiled terms alphabetically, dividing them into

the categories that integrated the preliminary systematic structure of the taxonomy; this was the third stage of the process.

Next, in a fourth stage, we prepared the first version of the taxonomic scheme, defining the hierarchical position of the terms within the respective categories and, consequently, ordering each specific term under the corresponding generic term. Subsequently, the terms were divided into two types of classes within each category: chains and rows. The former are vertical series of concepts and the latter are horizontal series of concepts; each series may be generic (type?) or partitive (part? whole?) (Campos e Gomes, 2007).

These procedures allowed us to move on to the fifth stage: preparing the classification superstructure of the taxonomy and filling it with all the terms. The final version includes the following 17 categories (see Fig. 4).

- 1) Administration
- 2) Fields of History
- 3) Historiographical sciences and techniques
- 4) Crown
- 5) Chronology

Figure 3. Formal and semantic normalization of the term "Cantigas de escárnio e de Maldizer"

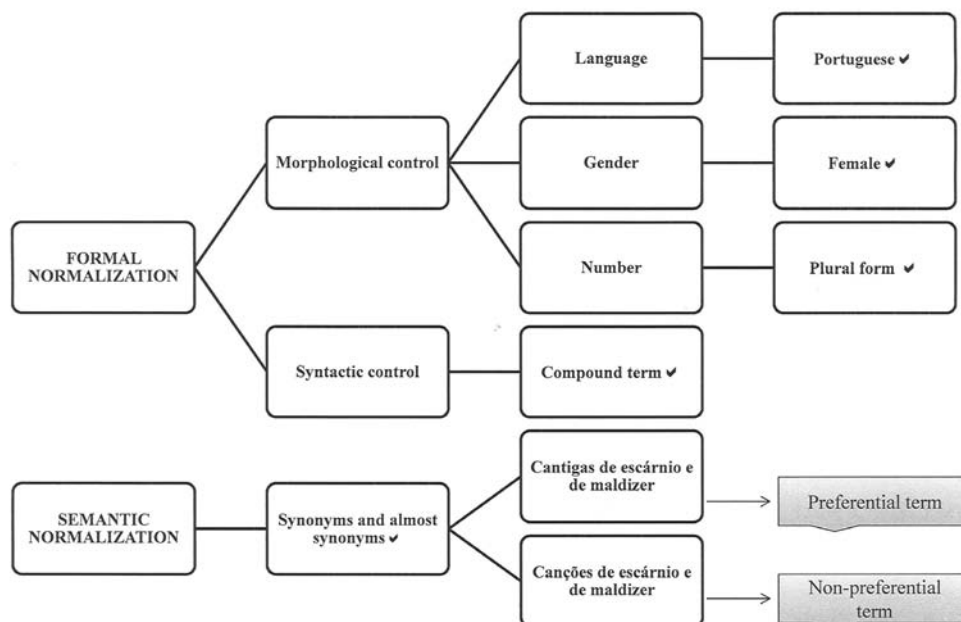


Figure 3: Formal and semantic normalization of the term “Cantigas de escárnio e de Maldizer”

- 6) Culture
- 7) Economy
- 8) Social structure
- 9) Historical sources
- 10) Geography
- 11) War
- 12) Personalities
- 13) Information resources
- 14) Reigns
- 15) Religion
- 16) Society
- 17) Private life

In a sixth stage, we evaluated the taxonomy according to the recommendations of the ANSI/NISO Z39.19-2005 standard, because both the NP 4036 and the ISO 25964-2011 standards fail to mention this issue. So, we chose to conduct a heuristic evaluation of the final version of the taxonomy, since we believe that this is the most appropriate approach to the controlled vocabulary we have built; it is a domain-based taxonomy and, as such, it implies a real and permanent intervention and monitoring of the construction process by experts, which is what actually happened (ANSI/NISO Z39.19, 2005).

Therefore, we asked for the formal collaboration of two experts, one from the field of Portuguese

Medieval History, and the other one from the field of Library and Information Science (LIS) –specialized in knowledge organization systems– whose evaluations we present further ahead.

In the seventh and last stage, we published the first version of the taxonomy. All the terms that had been compiled and normalized were fed into a specific taxonomy management software called Knowledge Manager (KM), marketed by the Spanish company *The Reuse Company*, which collaborates with the Departamento de Biblioteconomía y Documentación of the Universidad Carlos III de Madrid. The version we used is a downloadable test version (5.0.0). The KM allows the creation of different types of reports which can be exported in various formats (alphabetic, hierarchical or as a glossary), thus enabling terms to be fed into other applications. We decided to present the final version of the taxonomy alphabetically, insofar as, in our opinion, this option makes it possible to navigate through the semantic structure of the taxonomy, allowing an overall view of its categories and of their respective chains and rows as well.

Take, as an example, the search term “Demanda do Santo Graal”. Figure 5 shows the interface of the search term in the KM and some filtering options enabled by this *software*.

In figure 6 is displayed the result of the research carried out, according to a hierarchical presentation in which we can observe the existing subordination levels between the term “Demanda do Santo Graal” and all its generic terms.

Figure 4. Overview of all categories of the taxonomy in the KM software

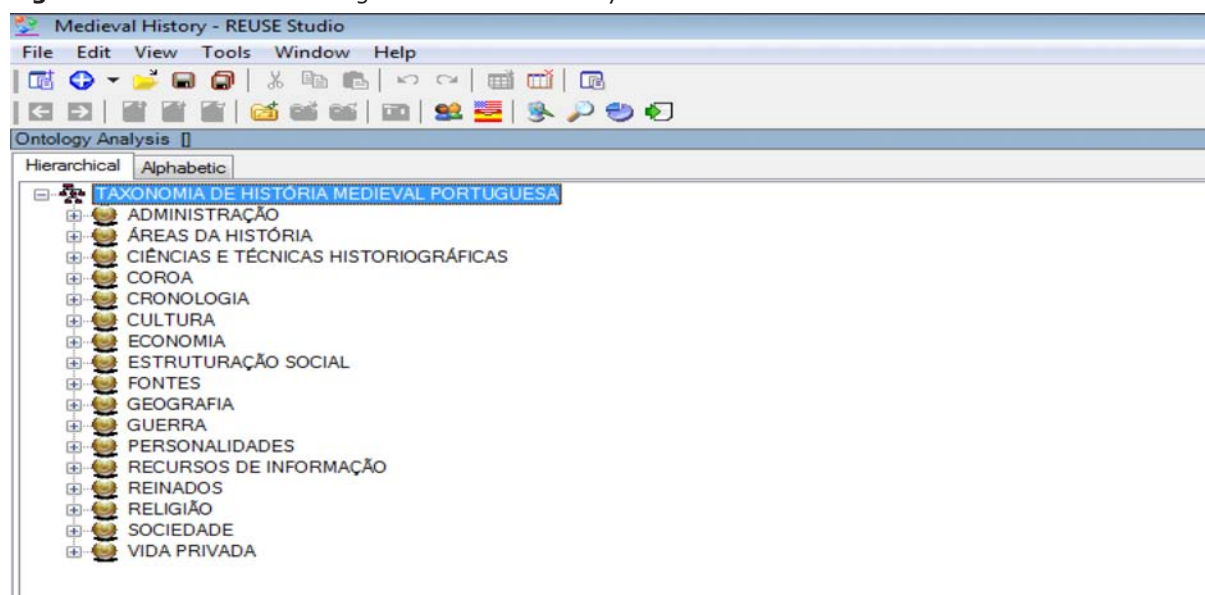


Figure 5. Search interface in the KM software (1)

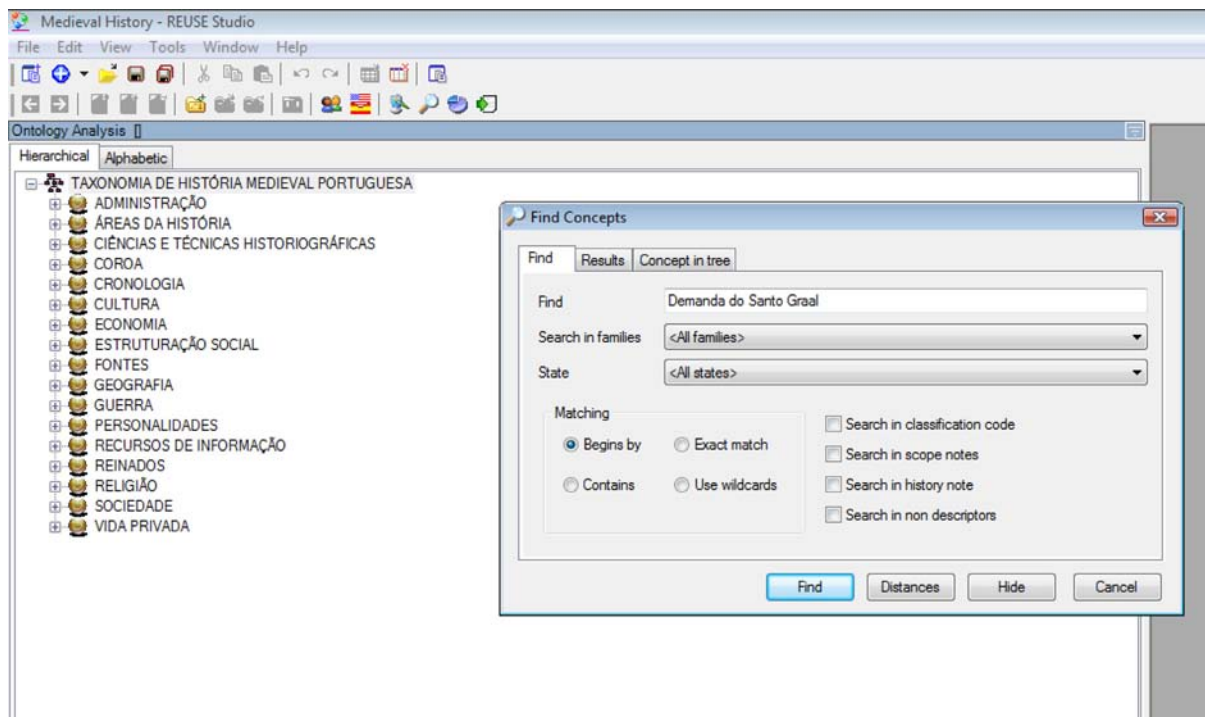
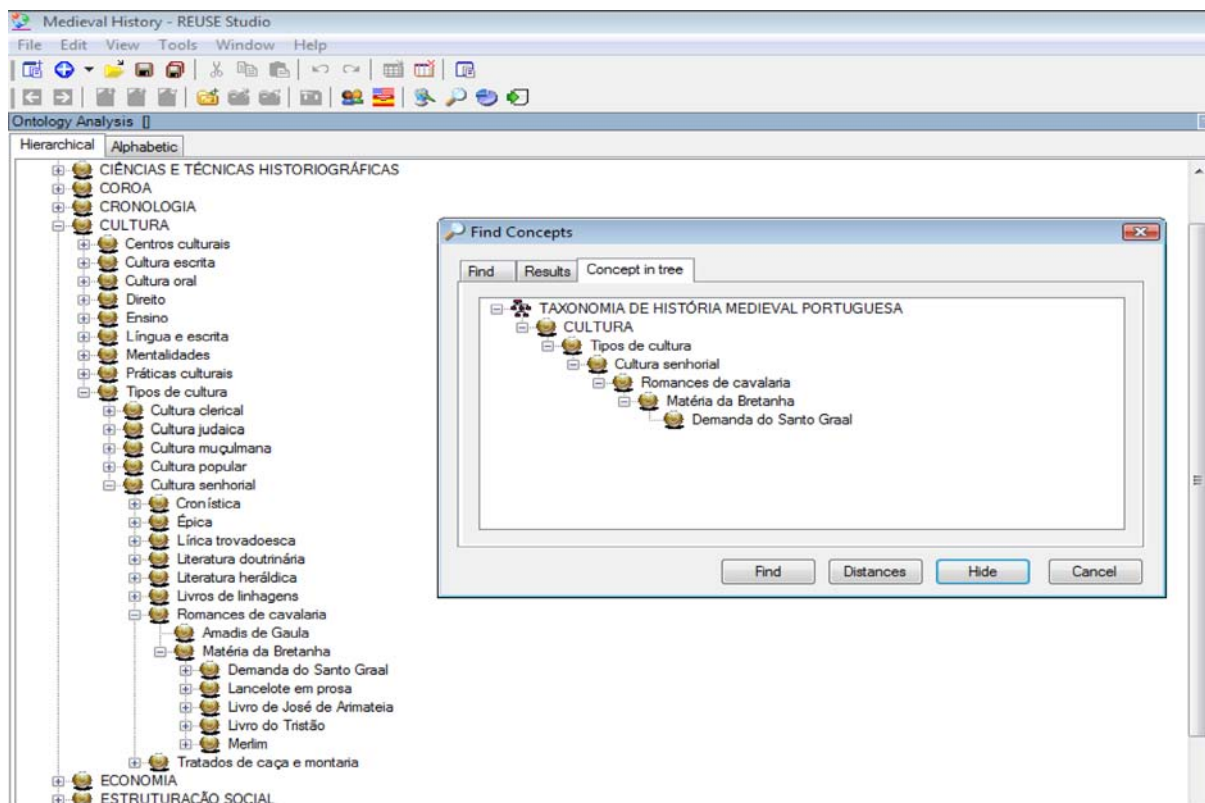


Figure 6. Search interface in the KM software (2)



3. PRESENTATION AND DISCUSSION OF THE RESULTS

Currently, the taxonomy includes 2987 terms which are constantly being reviewed and updated.

The full version of the taxonomy is available online at: <http://www.en.cidehus.uevora.pt/Bases-de-Dados/Taxonomia-de-Historia-Medieval-Portuguesa>

As shown in figure 7, the Portuguese Medieval History taxonomy has 17 categories, with different depths.

So, «Chronology» is the largest category, with 570 terms, while «Historiographical sciences and techniques» has just 9 terms.

The categories «Personalities», «Economy», «War», «Culture» and «Society» also stand out with 523, 486, 265, 244 and 231 terms, respectively.

On an intermediate level we list the categories «Administration», «Social structure», «Religion» and «Historical sources», which have, correspondingly, 142, 132, 114 and 93 terms.

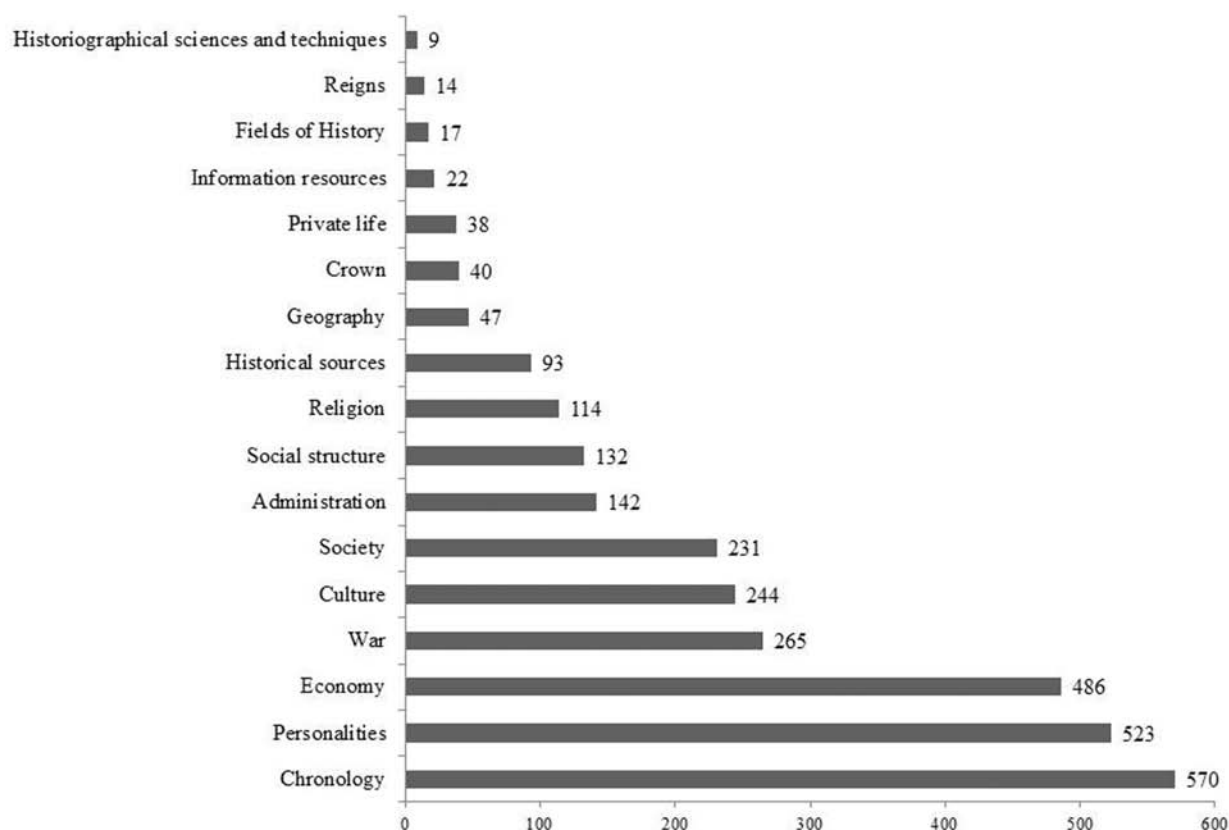
Finally, and with a smaller number of terms, we identify the categories «Geography», «Crown»,

«Private life», «Information resources», «Fields of History», «Reigns» and «Historiographical sciences and techniques», which have 47, 40, 38, 22, 17, 14 and 9 terms, respectively.

We have to consider the diverse sizes and depths of categories. This is due mainly to the different scopes of these specific categories. Other factors are the uneven distribution of specific sources for each category, and the variable extent to which different topics of Medieval History are represented in the scientific knowledge under evaluation - that which is developed within Portuguese universities.

With regard to the evaluation of the taxonomy, generally speaking, we asked our medievalist to conduct a semantic evaluation (regarding timeliness and communicability of the information) in addition, we asked our expert on representation of information to conduct a formal evaluation (on the morphological and syntactic aspects). For that purpose, we prepared two analysis grids, following the recommendations of the ANSI/NISO Z39.19-2005 standard. As expected in the case of interdisciplinary research works, the results of both evaluations were very positive and complemented

Figure 7. Distribution of the no. of terms across the categories of the taxonomy of Portuguese Medieval History



each other; this circumstance is essential for the work under study, given the need for a joint terminological validation.

Tables I and II show the evaluation criteria that were defined, as well as the score assigned by the experts for each criterion.

4. CONCLUSIONS

By way of conclusion, we make some essential considerations:

First, it should be noted that the taxonomy we constructed, when applied to research, successfully tested the methods previously tried by experts for designing, developing and maintaining controlled vocabularies of this nature.

We should also note that, although this taxonomy is representative of one country at one particular historical moment, it can still function as a starting point for building other controlled vocabularies with the same thematic scope, referring to different geographical locations, in the same time frame.

Finally, we stress the pioneering nature of this taxonomy within the panorama of Portuguese

Medieval Studies. Precisely due to the absence of comparative structures, both at the national and at the international levels, its construction was a great challenge that we were only able to meet thanks to an interdisciplinary collaboration between medievalists and I&D professionals. That is why we would like to stress the idea that the taxonomy we have built is, above all, a first attempt at defining a terminological categorization for this subject matter. As such, it is open for discussion and it undergoes a permanent formal and semantic evolution process, as the terms of the taxonomy are applied to existing specialized databases, on an experimental basis for the time being.

5. FUTURE DEVELOPMENTS

With regard to future developments, we consider that the taxonomy can integrate various materials from various digital resources (databases, bibliographies and other specialized compilations) into a single format (xml) and classify them all using the same taxonomy.

We would also like to emphasize that the reference taxonomy is still under construction because, even though the semantic structure is finished, we still

Table I. Evaluation grid (medievalist)

EVALUATION CRITERIA	SCORE
1. General and specific information sources	Very good
2. Uniformity with regard to the representation of the various subjects included in the taxonomy	Very good
3. Semantic relevance of the terms considering the vocabulary used by the users (timeliness of the terms)	Very good
4. Adequacy of the hierarchical structure	Very good
5. Efficiency of the browsing system	Very good
6. Efficiency of the search system (retrieval and search options)	Very good
7. Notes/Suggestions: The researchers should ask for more specific advice from experts related to the various thematic fields covered by the taxonomy.	

Table II. Evaluation grid (expert on representation of information)

EVALUATION CRITERIA	SCORE
1. Reference sources and procedures used in the normalization of the vocabulary	Very good
2. Morphological and syntactic consistency of the terms	Very good
3. Conceptual accuracy in terms of the definition of the hierarchical relationships	Very good
4. Efficiency of the browsing system	Very good
5. Efficiency of the search system (retrieval and search options)	Good
6. Notes/Suggestions: The researchers should aim at diversifying and refining the available search options.	

have to develop a way to determine names through labels and schematic links which favours its reuse (Daconta et al, 2003). So, this is an essentially terminological taxonomy that does not have an ontological nature. In a second stage, we should define a naming convention aimed at enabling unambiguous identifications through the use of a formalized representation format, whose patterns should be determined according to a proper standard. This will, therefore, be the main future challenge for the taxonomy we have presented here.

7. REFERENCES

Bibliography on medieval history

Note: Since it is impossible to include all the bibliographic references that were used as sources of information for the selection of the terms included in the taxonomy, we decided to mention only general histories and reference university textbooks from the field of Portuguese Medieval Studies.

Branco, M. J. V.; Costa, A. M. (1992). *História da Idade Média: caderno de apoio*. Lisboa: Universidade Aberta. ISBN 972-674-370-2.

Coelho, M. H.da C. (1991). *História medieval de Portugal: guia de estudo*. Porto: Universidade Portucalense. ISBN 972-9354-07-3.

Marques, A. H. De O. (1988). *Guia do estudante de História medieval portuguesa*. 3.^a ed. Lisboa: Estampa.

Mattoso, J. (dir.) (1997). *História de Portugal: a monarquia feudal (1096-1480)*. Lisboa: Estampa, vol. 2. ISBN 972-33-1263-8.

Moreno, H. B. (1995). *História de Portugal medievo: político e institucional*. Colab. Maria da Conceição Falcão Ferreira, Luís Carlos do Amaral, Luís Miguel Duarte. Lisboa: Universidade Aberta, 2 vols. ISBN 972-674-132-7 (vol. 1). ISBN 972-674-135-1 (vol. 2).

Ramos, R.; Sousa, B.V.; Monteiro, N. G. (2009). *História de Portugal*. Lisboa: A Esfera dos Livros. ISBN 978-989-626-139-9.

Serrão, J. V. (1994). *História de Portugal: Estado, pátria, nação*. 5.^a ed. Lisboa: Verbo, vol. 1. ISBN 972-22-0271-5.

Serrão, J.; Marques, A. H. de O. (dir.) (1987). *Nova História de Portugal: Portugal na crise dos séculos XIV e XV*. Coord. A. H. de Oliveira Marques. Lisboa: Estampa, vol. 4.

Serrão, J.; Marques, A. H. de O. (dir.) (1996). *Nova História de Portugal: Portugal em definição de*

6. ACKNOWLEDGEMENTS

We would like to thank the medievalists and the I&D professionals for their collaboration in the semantic and formal validation of the taxonomy, respectively.

This article was also made possible thanks to the collaboration of the Spanish company *The Reuse Company*, which granted us free access to the KM *software* that allowed us to build the taxonomy and publish its first version.

fronteiras (1096-1325): do Condado Portucalense à crise do século XIV. Coord. Maria Helena da Cruz Coelho, Armando Luís de Carvalho Homem. Lisboa: Presença, 1996. ISBN 972-23-2039-4, vol. 3.

Tavares, M. J. F. (1990). *Sociedades e culturas portuguesas*. Lisboa: Universidade Aberta, 2 vols. ISBN 972-674-030-4 (vol. 1). ISBN 972-674-049-5 (vol. 2).

Tavares, M. J. F. (1992). *História de Portugal medievo: economia e sociedade*. Lisboa: Universidade Aberta. ISBN 972-674-100-9.

Bibliography on representation of information

ANSI/NISO Z39.19-2005. *Guidelines for the construction, format and management of monolingual controlled vocabularies*. NISO Press: Bethesda, MD. ISBN 1-880-124-65-3.

Campos, M. L.; Gomes, H. Espanha a. Taxonomia e classificação: a categorização como princípio. *Enancib: Encontro Nacional de Pesquisa em Ciência da Informação*. [Electronic Version]. Salvador, 8, 2007. Retrieved 22 sep. 2010, from <http://www.enancib.ppgci.ufba.br/artigos/GT2--101.pdf>

Conway, S.; Sligar, C. (2002). *Unlocking knowledge assets*. Redmont: Microsoft Press. ISBN 0-7356-1463-6.

Cumming, M. (2003). Tomatoes are not the only fruit: a rough guide to taxonomies, thesauri, ontologies and the like. *Record Management Bulletin*, nº 113, 1-4.

Daconta, M.; Obrst, L. J.; Smith, K. T. (2003). *The semantic web. A guide to the future of XML, web services and knowledge management*. Indianapolis: Wiley.

Gilchrist, A. (2003). Thesauri, taxonomies and ontologies: an etymological note. *Journal of documentation*, vol. 59, (1) 7-18. ISSN 0022-0418. <http://dx.doi.org/10.1108/00220410310457984>

- ISO 25964-2011. Information and Documentation. *Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval. Part 2: Interoperability with other vocabularies*. Geneva: International Organization for Standardization.
- Jagerman, E. (2006). *Creating, maintaining and applying taxonomies*. Zoetermeer: E. Jagerman. ISBN 90-811-274-1-1.
- Medeiros, F. (2014). *A historiografia medieval portuguesa na viragem do milénio: análise bibliométrica (2000-2010) e representação taxonómica* [Electronic Version]. Retrieved 22 oct. 2014, from <https://dspace.uevora.pt/rdpc/handle/10174/11220?mode=full>
- Moreiro González, J. A. (2006). La representación y recuperación de los contenidos digitales: de los tesauros conceptuales a las folksonomías. In: Tramullas, J. (coord.). *Tendencias en Documentación Digital*. Gijón: TREA: 81-109.
- Moreiro González, José Antonio (2011). *Linguagens documentárias e vocabulários semânticos para a web: elementos conceituais*. Salvador: EDUFBA. ISBN 978-85-232-0824-0.
- NP 4036 (1992). Documentação. *Tesauros monolíngues: directivas para a sua construção e desenvolvimento*. Lisboa: Instituto Português da Qualidade.
- Portugal. Biblioteca Nacional. Área de Indexação e Classificação (1998). *SIPORBASE: sistema de indexação em português: manual*. 3.ª ed. rev. e aum. Lisboa: Biblioteca Nacional. ISBN 972-565-154-5.
- Zhonghong, W.; Chaudhry, A. S.; Khoo, C. (2006). Potential and prospects of taxonomies for content organization. *Knowledge organization: international journal devoted to concept theory, classification, indexing and knowledge representation*, vol. 33 (3), 160-170. ISSN 0943-7444.



CRÍTICA DE LIBROS / BOOK REVIEWS

Gephi Cookbook

Devangana Khokhar

Birmingham, Packt publishing, 2015. 274 pags. ISBN 978-1-78398-740-5

Gephi y su contexto

El 31 de Julio de 2008, Mathieu Bastian lanzó por primera vez Gephi, una plataforma para la visualización interactiva y la exploración de todo tipo de redes, sistemas complejos y grafos dinámicos y jerárquicos. Junto a sus colaboradores Sébastien Heymann y Mathieu Jacomy la habían desarrollado para "importar, exportar, manipular, analizar, filtrar, representar, detectar comunidades y exportar grandes grafos y redes" (Bastian y otros, 2009). El propio Sébastien Heymann (2014) ofrece una descripción más actual y completa de la aplicación, cuya versión 0.9 se lanzó el pasado 21 de diciembre.

Gephi surgió en un panorama claramente dominado por aplicaciones mono-plataforma, parciales, desequilibradas y con una curva de aprendizaje realmente abrupta. Dos muy populares, Pajek y Ucinet/Netdraw, están limitadas al entorno Windows y se inclinan hacia el análisis de grafos (Pajek) o hacia la visualización (Ucinet/Netdraw, GUESS). NodeXL es la abreviatura de Network Overview Discovery Exploration for Excel; NetworkX es una biblioteca de Python e iGraph un módulo de R, mientras que NetLogo es un entorno de modelado. Entre las restantes aplicaciones destacan las destinadas a la exploración y visualización de redes de datos bibliométricos: HistCite, SciMAT (Cobo y otros, 2012) VOSviewer y CitNetExplorer (van Eck, 2014) CiteSpace (Chen, 2014) o Science of Science (SCI2) tool. Con la aparición de Bibliotools, achacar a Gephi la falta de una funcionalidad específica para el procesamiento de datos bibliométricos (van Eck 2014) ha dejado de tener sentido. Gephi es muy útil y cada vez más utilizado en el análisis y visualización de redes sociales y en bibliometría. De su popularidad dan cuenta algunos datos sencillos: de las 246 descargas de la primera versión

se ha pasado a un total de 1.272.293 descargas en septiembre de 2015. A pesar de que aún está en fase beta, ha recibido más de 630 citas en el WoS; además, 877 de los documentos de Scopus contienen el término "Gephi" en sus referencias bibliográficas. En palabras de Chaomei Chen (2014) "Gephi has become competitive even to the most mature and widely used software available from the earlier generations such as Pajek".

En España, tanto el grupo Scimago como el grupo SCI2S han producido interfaces gráficas propias para la visualización de datos bibliométricos, pero son excepciones en un panorama muy modesto. Casi 200 trabajos de investigación españoles incluidos en Scopus citan (han utilizado) Pajek; de ellos, 116 se han publicado desde 2011. La misma fuente ofrece 48 trabajos con contribución española publicados desde 2011 que citan Gephi. La rutina (García et al. 2015) y el recurso a aplicaciones muy especializadas para proyectos concretos (Nualart-Vilaplana y otros, 2014) pueden ser causas de ese desequilibrio pero, cuando se trata de un conjunto de aplicaciones, de una nueva plataforma en continuo desarrollo. Hay otra razón de peso: la falta de información y de formación.

La webografía y la bibliografía sobre Gephi

El problema con Gephi, el mismo que afecta a otras aplicaciones y plataformas colaborativas, se ha puesto de manifiesto de forma explícita (Rees, 2010): "The only thing keeping Gephi from being a professional grade tool is its bugginess and lack of documentation. If its development community can sustain it, Gephi should be able to improve and become a more robust software tool". En efecto, la naturaleza abierta puede resultar una desventaja para algunas aplicaciones en relación a su docu-

mentación y a su soporte. El grupo central de desarrolladores de Gephi y la comunidad de usuarios más activos, a pesar de su continuada producción de tutoriales, no bastan para abastecer a la muy heterogénea comunidad de usuarios y sus interrogantes, a pesar de que es precisamente la interacción entre unos y otros lo que favorece la continua producción de módulos y complementos, propios de una aplicación tan extensible.

Hasta hace poco, el conocimiento sobre el uso de Gephi dependía más de presentaciones y videos (Levallois 2014) entradas en blogs (Hirst 2010; Jacomy 2015) y artículos en wikis (<https://github.com/gephi/gephi/wiki>) que de uno o más textos "canónicos". Ante la falta de un manual oficial, se necesitaban obras que ofrecieran una descripción de las características de Gephi y un buen relato de su funcionamiento. Y esas obras llegaron.

El primer libro comercial sobre Gephi apareció en Septiembre de 2013 (Cherven, 2013). A pesar de escribir con fluidez y cierta espontaneidad, su autor, un analista de mercados de General Motors, abusaba de la llamada a la experimentación, al "prueba tu mismo". Este recurso ocultaba la falta de sistematización, de un plan para el desarrollo de la obra que, así, contenía muchas más omisiones que aciertos y parecía picotear nerviosamente aquí y allá, a lo largo de la interfaz de la aplicación (Amat, 2014). Quizá se trataba de una primera toma de contacto puesto que el mismo autor anunció, y finalmente ha publicado en enero de este mismo año, una nueva entrega en apariencia más elaborada. Casi coincidiendo con el segundo libro de Cherven, ha aparecido el "recetario" sobre Gephi. Su autora, Devangana Khokhar, es una ingeniera informática tan amiga de los grandes números (big data) como de las grandes distancias: es consultora de ThoughtWorks en Illinois y miembro del capítulo en Bangalore de Women Who Code. Su experiencia es palpable a lo largo de todo el libro.

Las recetas con Gephi

Los recetarios sobre aplicaciones informáticas se han popularizado. Su organización es casi invariable: una serie de unidades de estructura homogénea que representan operaciones y se agrupan por su finalidad en secciones o capítulos. En el caso de este recetario con Gephi, comprenden:

- Los preparativos (getting ready, los ingredientes)
- Las operaciones (how to do it..., la receta propiamente dicha)
- Los detalles (how it works, la descripción de los ingredientes)

- Para saber más (there's more..., variaciones y operaciones adicionales)
- Ver también (see also, reenvíos a otros capítulos, webliografía y referencias a otros materiales).

Las unidades se organizan después en capítulos que definen operaciones, desde el inicial sobre la instalación y la descripción de la interfaz básica del programa hasta el octavo, dedicado a un repaso de las redes dinámicas y las redes jerárquicas, con el añadido de dos capítulos finales dedicados a los conjuntos de datos ya elaborados y las extensiones de la aplicación. Los repaso a continuación.

Lo que en el libro de Cherven era un capítulo inicial incompleto, con olvido de elementos centrales en la interfaz (Amat, 2014) aquí se resuelve de manera muy ordenada y lógica, con la presentación de la red que va a acaparar la exploración de casi todas las operaciones que el libro presenta, la popular red que conecta los personajes de "Los Miserables". A partir de su carga, se describen someramente el laboratorio de datos, la pantalla de previsualización y la interfaz de trabajo (llamada vista general en la versión española) con sus módulos asociados.

El segundo capítulo, que contiene 14 recetas, se inicia con una sobre la generación de grafos aleatorios que parece estar fuera de lugar, y cuya utilidad se despacha con referencias al modelo original de Erdős y Rényi y a su revisión por Newman (2002). Salvando esta primera unidad, el capítulo ofrece un repaso muy completo de las posibilidades que Gephi ofrece para la exploración visual de una red, sus vértices (nodos) y sus aristas (o arcos, conexiones). Las recetas aparecen organizadas y en un orden lógico que responde al título de esta primera serie de operaciones, "Basic Graph Manipulations".

La representación gráfica y el análisis estadístico

Los capítulos del tres al cinco presentan las demás operaciones que Gephi puede realizar sobre las redes que explora. La presentación de los algoritmos para la representación de las redes, en el capítulo tres, no se limita a relacionar los principales. Además, la autora ofrece la suficiente información adicional en cada apartado "there's more" como para entender el mecanismo de representación. Hay un exceso de detalle en algunos casos; por ejemplo, cuando se dedican recetas diferentes al algoritmo de rotación en el sentido de las agujas del reloj y al correspondiente a la rotación en sentido inverso, al igual que también se separan los algoritmos de contracción y expansión de las gráficas. Y al contrario, no se menciona OpenOrd (Martin y otros, 2011) uno de los algoritmos más rápidos desarrollados e incorporados a Gephi re-

cientemente. Destaca, eso sí, en este capítulo las referencias externas a las descripciones completas de los algoritmos de representación.

El cuarto capítulo quizá se debería haber aplazado hasta después de presentar las opciones de análisis cuantitativo de las redes. Como está dedicado a la exploración estructural y ésta se basa en la representación de los atributos de nodos y aristas, además de su agrupamiento, se recurre a atributos cuantitativos (grado, modularidad) que se anticipan al capítulo siguiente. En todo caso, la red de protagonistas de *Los Miserables*, que tiene formato gexf (lo comento más adelante) ya incorpora como atributo la clase de cada nodo y algunas de sus propiedades de visualización como el tamaño, el color y la posición. También el valor de cada enlace. El capítulo presenta las posibilidades de manipulación de los colores y tamaños (o grosores) de nodos y aristas, así como el etiquetado de unos y otras en función de sus atributos.

La métrica de las redes se trata a lo largo del capítulo cinco. Las propiedades estadísticas de los nodos, las aristas y la red en su conjunto se presentan en 14 recetas que, si bien no superan el nivel descriptivo de las anteriores, ofrecen con sencillez las definiciones de algunos de los parámetros que se presentan. El capítulo contiene algunas imprecisiones sobre el PageRank, algunos descuidos sobre el cálculo de la modularidad o del valor de HITS y concede poca relevancia al coeficiente de agrupamiento (clustering coefficient, transitivity) pero es el más instructivo de toda la obra. Y ello a pesar de que no explora el hecho de que cualquier cálculo sobre nodos y aristas se traduce automáticamente en la aparición de nuevas columnas en el laboratorio (tabla) de datos. Y tampoco se refiere a que cualquier filtro de la red también se refleja allí.

Estos efectos se podrían haber comentado en el sexto capítulo, "Working in the Data Laboratory Mode" pero la autora lo estructura con una receta introductoria sobre la importación de datos de red en formato tabular seguida de otras 10 que repasan las operaciones de manipulación de nodos y relaciones y sus atributos. Apenas se comenta las repercusiones que estas operaciones tienen en la topología de la red, vista a través de la interfaz de trabajo. De este modo, se pierde la posibilidad de presentar las opciones de importación de archivos de red, que son muy flexibles en Gephi y justifican su interoperabilidad con otras aplicaciones.

El capítulo dedicado a la visualización previa de las redes trazadas es un mero repaso a los diseños preinstalados en la aplicación y ni siquiera contiene una descripción de los controles para la preparación de la imagen.

Las redes dinámicas y el formato gexf

En el octavo capítulo, se tratan las redes dinámicas en Gephi, pero con demasiada brevedad y con excesiva simpleza. Utilizando el generador de redes dinámicas de Gephi, se exploran su visualización animada y la representación gráfica de algunos valores y atributos en las celdas correspondientes del laboratorio de datos. Pero no se explora el proceso para generar e importar (o exportar) redes dinámicas ni tampoco la estructura del formato que permite generarlas. Una de las mayores fortalezas de Gephi es el Graph Exchange XML Format (gexf) que sus mismos creadores desarrollaron a partir de 2007. La plasticidad de este formato es extraordinaria, soporta tanto las redes estáticas como las dinámicas y sirve de soporte al intercambio con otros programas (no con Pajek) al tiempo que permite el diseño de una red con el mismo procesador de textos con el que se crea cualquier archivo XML. Habida cuenta del auge que experimenta el análisis de redes dinámicas, tanto las redes como el formato que las describe hubieran merecido una presentación más completa, como la que les dedica en su segunda obra Cherven (2015).

Notas finales

Este Gephi Cookbook se define como "una guía para el aprendizaje de la exploración y visualización interactivas de redes, acompañada de los conceptos de teoría de grafos que las soportan". Y se dirige a "aquellos que, sin experiencia previa, desean aprender tanto esas operaciones como los conceptos subyacentes".

La obra cumple con creces esos objetivos y es recomendable como un manual de introducción al uso de Gephi. Su lectura, combinada con la de una obra de divulgación como la breve introducción a las redes de Guido Caldarelli y Michele Catanzaro (2012) puede favorecer una toma de contacto realmente útil con el dominio del análisis de redes, sea cual sea su campo de aplicación.

Cualquier análisis de redes comienza con datos y esos datos se encuentran en soportes con diferentes formatos, sean tablas, bases de datos o meros ficheros planos. Incorporar esos datos a Gephi es la tarea inicial en cualquier proyecto que recurra a esa magnífica aplicación. Por esta razón se hace necesario completar la lectura de las recetas con el seguimiento de otros materiales, por ejemplo la presentación "Gephi focus on data import" (Levallois, 2014). Del mismo modo, para obtener una idea sobre el uso y análisis de tablas dinámicas es conveniente recurrir al reciente texto de Cherven (2015) o al tutorial de Eduardo Ramos (el desarrollador zaragozano de Gephi) sobre importación de datos dinámicos (Ramos, 2015).

Todos los recetarios, hasta los excelentes ejemplares de las colecciones de O'Reilly, se ven consustanciados por sus propios esquemas y nunca resultan totalmente exhaustivos. Los olvidos de Devangana Khokhar son comprensibles. Su libro es recomendable y Gephi merece su lectura.

Carlos Benito Amat

Ingenio (CSIC-UPV)

REFERENCIAS

- Amat, C. B. (2014). Reseña de Network Graph Analysis and Visualization with Gephi by Ken Cherven. *Redes: revista hispana para el análisis de redes sociales*, 25(1), 201-209. <http://dx.doi.org/10.5565/rev/redes.499>
- Bastian, M.; Heymann, S.; Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. San Jose, Third International AAAI Conference on Weblogs and Social Media, 8, 361-362. <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/viewFile/154/1009>
- Caldarelli, G.; Catanzaro, M. (2012). *Networks: a very short introduction*. Oxford University Press. La versión española (*Redes: Una breve introducción*. Madrid, Alianza) apareció en 2014. <http://dx.doi.org/10.1093/actrade/9780199588077.001.0001>
- Chen, C. (2014). *Mapping Scientific Frontiers: The Quest for Knowledge Visualization*, 2nd ed. New York: Springer. <http://dx.doi.org/10.1007/978-1-4471-5128-9>
- Cherven, K. (2013). *Network graph analysis and visualization with Gephi*. Birmingham, Packt Publishing Ltd.
- Cherven, K. (2015). *Mastering Gephi Network Visualization*. Birmingham, Packt Publishing Ltd.
- Cobo, M. J.; López-Herrera, A. G.; Herrera-Viedma, E.; Herrera, F. (2012). SciMAT: A new science mapping analysis software tool. *Journal of the American Society for Information Science and Technology*, 63(8), 1609-1630. <http://dx.doi.org/10.1002/asi.22688>
- García, A. G.; Ibáñez, A. P.; Sapena, A. F.; Mancebo, M. F. P.; Moreno, L. M. G. (2015). Herramientas de análisis de datos bibliográficos y construcción de mapas de conocimiento: Bibexcel y Pajek. *BiD: Textos universitaris de biblioteconomia i documentació*, (34), 11. <http://dx.doi.org/10.1344/BiD2015.34.22>
- Levallois, C. (2014): Gephi Tutorials and Plugins. <http://www.clementlevallois.net/gephi.html> [consulta 15-9-2015].
- Heymann, S. (2014): Gephi. En Alhajj, R.; Rokne, J. (Eds.) *Encyclopedia of social network analysis and mining*. Springer. pp. 612-625. http://dx.doi.org/10.1007/978-1-4614-6170-8_299
- Hirst, T. (2010). Getting Started With The Gephi Network Visualisation App – My Facebook Network, Part I. Recuperado a partir de <http://blog.ouseful.info/2010/04/16/getting-started-with-gephi-network-visualisation-app-my-facebook-network-part-i/>
- Jacomy, M. (2015). Gephi blog. <https://gephi.wordpress.com/> [consulta 10-9-2015]
- Martin, S; Brown, W. M.; Klavans, R.; Boyack, K. W. (2011). OpenOrd: An Open-Source Toolbox for Large Graph Layout. En P. C. Wong y otros. (Eds.), *Proc. SPIE 7868: Visualization and Data Analysis*, p. 786806). San Francisco. <http://doi.org/10.1117/12.871402>
- Newman, M. E. (2002). Random graphs as models of networks. arXiv preprint <http://arxiv.org/abs/cond-mat/0202208v1> / <http://dx.doi.org/10.1002/3527602755.ch2>
- Nualart-Vilaplana, J.; Pérez-Montoro, M.; Whitelaw, M. (2014). Cómo dibujamos textos: Revisión de propuestas de visualización y exploración textual. *El profesional de la información*, 23(3), 221-235. http://www.elprofesionaldelainformacion.com/contenidos/2014/mayo/02_esp.pdf / <http://dx.doi.org/10.3145/epi.2014.may.02>
- Ramos, E. Import Dynamic Data. <https://github.com/gephi/gephi/wiki/Import-Dynamic-Data> [consulta 16-9-2015].
- Rees, K. (2010). Data Visualization Review: Gephi, Free Graph Exploration Software. Information Aesthetics. http://infosthetics.com/archives/2010/07/review_gephi_graph_exploration_software.html. [Consulta 10-9-2015].
- van Eck, N. J.; Waltman, L. (2014). Visualizing bibliometric networks. En Ding, Y.; Rousseau, R.; Wolfram, D. (Eds.) *Measuring Scholarly Impact: Methods and Practice*. New York: Springer. pp 285-320. http://dx.doi.org/10.1007/978-3-319-10377-8_13



NOTICIAS / NEWS

El Centro de Ciencias Humanas y Sociales del CSIC galardonado con el Premio MEDES 2015 en la categoría de mejor institución por su labor de coordinación y mantenimiento del Sistema LATINDEX en España

El pasado 2 de noviembre de 2015 tuvo lugar el acto de entrega de los Premios MEDES-MEDICINA en Español otorgados por la Fundación Lilly con el objeto de reconocer las iniciativas que promueven nuestro idioma en la comunicación del conocimiento de las ciencias de la salud. Con la lectura del dictamen del jurado se anunció que el galardón, en la categoría de Mejor institución, había recaído en el Centro de Ciencias Humanas y Sociales del CSIC (CCHS) con motivo de su trabajo de coordinación y mantenimiento de la base de datos online del Sistema LATINDEX en España.

El director de la Fundación Lilly, José Antonio Sacristán, afirmó que esta base de datos online "contribuye desde 1999 de manera relevante en beneficio de la producción y publicación científica biomédica en español, promoviendo con ello el uso del idioma español para la transmisión del conocimiento científico, en general, y biomédico, en particular" así "Cabe resaltar de este proyecto su capacidad para favorecer la visibilidad de las revistas españolas, siendo muy valorado su elevado grado de actualización, lo que le ha convertido en un recurso de referencia en las universidades españolas y americanas".

El director de la Fundación aseguró que "en línea con el compromiso de la Fundación Lilly de fomentar el uso del español como lengua de difusión del conocimiento científico, con estos galardones queremos contribuir a mejorar el uso del lenguaje médico en español, y concienciar sobre la importancia de utilizar de forma apropiada nuestro idioma".

LATINDEX es el acrónimo de Sistema Regional de Información en Línea para Revistas Científicas

de América Latina, el Caribe, España y Portugal (www.latindex.org). El proyecto LATINDEX abarca un amplio espectro de disciplinas científicas, entre las que la Medicina tiene gran relevancia, y tiene como objetivo principal difundir, hacer accesible y elevar la calidad las revistas académicas, a través de la cooperación.

El sistema se centra en la comunicación de la investigación realizada en el ámbito iberoamericano, apoyando la profesionalización del trabajo editorial e impulsando la calidad editorial de las revistas científicas y técnico-profesionales. En esta tarea, el proyecto LATINDEX –impulsado y coordinado por la Universidad Nacional Autónoma de México– está acompañado casi desde sus inicios por el Consejo Superior de Investigaciones Científicas, socio institucional de LATINDEX en España desde el año 1998.

El contacto inicial del CSIC con LATINDEX tuvo lugar en la tercera reunión técnica del Sistema a finales de 1997, solo unos meses después de la reunión fundacional del proyecto. En febrero de 1998 se oficializó la representación española a través del Centro de Información y Documentación Científica del CSIC (CINDOC), que fue a partir de entonces el centro de acopio LATINDEX para las revistas del país.

En 2007 se creó el Centro de Ciencias Humanas y Sociales (CCHS) del CSIC (www.cchs.csic.es), y el CINDOC se integró en él pasando a denominarse Instituto de Estudios Documentales sobre Ciencia y Tecnología (IEDCYT). La colaboración con LATINDEX continuó en este nuevo marco organizativo y se creó dentro de este instituto el Grupo de Investigación "Evaluación de las publicaciones científicas" (EPUC)

que fue, hasta 2013 y en colaboración con la Unidad de Bases de Datos ICYT, el referente de LATINDEX en España. Actualmente el Centro de Ciencias Humanas y Sociales del CSIC sigue siendo el socio español del sistema y la Unidad de Bases de Datos ISOC e ICYT actúa como centro nacional de acopio de información para las revistas del país: califica las publicaciones académicas para la entrada en el Catálogo y asesora a los editores españoles sobre los criterios de calidad editorial. En la relación con los coordinadores de Latindex en México, y en el desarrollo y difusión de las acciones de Latindex, el Centro de Ciencias Humanas y Sociales del CSIC está apoyado también por Editorial CSIC y la Vicepresidencia de Organización y Relaciones Institucionales del CSIC.

Gloria Fernández-Mayoralas, directora del Centro de Ciencias Humanas y Sociales (CCHS), agradeció el premio y quiso compartir especialmente el honor concedido por la Fundación con la presidenta de Latindex, el coordinador general, y todos los socios iberoamericanos que representan a 22 países participantes y contribuyen a desarrollar las acciones de Latindex sobre la base de la cooperación.

La Fundación Lilly, que el año 2015 ha cumplido su decimoquinto aniversario, tiene como objetivo contribuir al mejor desarrollo de la sanidad española favoreciendo la generación y divulgación del conocimiento al servicio de los ciudadanos. Sus actividades se orientan al apoyo y fomento de la investigación (Premios de Investigación Biomédica o Foro de Ciencia); la divulgación del conocimiento (MEDES – MEDicina en ESpañol, reuniones y encuentros de alto nivel científico); y la formación y promoción de los profesionales sanitarios (Programa IESE, Cátedra de Educación Médica). En sus iniciativas se tratan temas novedosos de marcada actualidad, contando para ello con la participación de personalidades sanitarias, científicos e investigadores de prestigio

contrastado y el aval del Consejo Científico Asesor de la Fundación.

La Fundación Lilly y el sistema LATINDEX comparten algunos objetivos. MEDES-MEDicina en ESpañol es una iniciativa de la Fundación Lilly que tiene como objetivo contribuir a promover la publicación científica en revistas biomédicas en español, así como favorecer su difusión nacional e internacional mediante diversas actividades de formación y divulgación. Dispone de una base de datos de revistas biomédicas en español, a las que se accede de forma gratuita a través de un potente buscador en Internet (www.medes.com). Esta herramienta de consulta bibliográfica a disposición del profesional sanitario actualiza y evalúa sus contenidos de forma continua. Por otra parte, las jornadas MEDES, que se celebran anualmente desde el año 2005, son un foro profesional fundamental para conocer la situación de las publicaciones biomédicas en España. En ellas se han analizado a fondo la calidad, la visibilidad y la proyección de las publicaciones, aspectos relevantes para las revistas de todas las áreas científicas, y en los que se centran las políticas y acciones que lleva a cabo LATINDEX.

Otra categoría de los premios MEDES-MEDicina en ESpañol, en concreto a la 'Mejor Iniciativa', tuvo como agraciada a la web Familia y Salud. Padres y pediatras al cuidado de la infancia y adolescencia (www.familiaysalud.es), por su originalidad y carácter participativo de todos los actores implicados en la salud y el enfoque basado en la evidencia. Sendas menciones especiales recayeron en DianaSalud.com y la Fundación Iberoamericana Down21 por ser finalistas a las categorías de 'Mejor Iniciativa' y 'Mejor Institución'.

Teresa Abejón Peña

Centro de Ciencias Humanas y Sociales del CSIC