

Estructura de la colaboración científica española en Biblioteconomía y Documentación (Scopus 1999-2007)*

Benjamín Vargas-Quesada**
David Minguillo***
Zaida Chinchilla-Rodríguez****
Félix de Moya-Anegón*****

Resumen

En este trabajo se identifica, visualiza y analiza la estructura principal de la colaboración científica de los autores en el área de Biblioteconomía y Documentación en España, a partir de los documentos registrados en la base de datos Scopus, entre los años 1999 y 2007. Con el objetivo de estudiar las comunidades de colaboración (subredes) que integran los componentes más relevantes en este dominio, se aplican medidas de similitud, técnicas de análisis de redes sociales, como la cercanía global y K-core, procedimientos bibliométricos, como el análisis de coocurrencia de autores y el análisis de frecuencia de términos. De esta forma, se logran identificar 21 componentes conformados por 28 subredes que revelan la micro-estructura de la red de colaboración científica. En la red se observa una colaboración centralizada en unos pocos componentes, originada en el sector investigador y que tiene como frentes de colaboración la recuperación de información y gestión de la información. Se destaca la escasa cooperación entre las subredes, donde tiene un papel importante la proximidad territorial e institucional, y la producción y evaluación científica son los frentes de colaboración más comunes. La metodología aplicada logra proporcionar una imagen general de la estructura principal de colaboración científica por autores en España, durante el periodo estudiado.

Palabras clave: colaboración científica, coautoría, redes de colaboración, análisis de redes sociales, centralidad global, biblioteconomía y documentación, España

* Artículo derivado de los proyectos de investigación realizados por el Grupo SCImago, Unidad Asociada del CSIC, en colaboración con Statistical Cybermetrics Research Group, School of Computing and Information Technology, University of Wolverhampton.

** Doctor en Documentación e Información Científica. Miembro del Grupo SCImago. Profesor Titular de la Facultad de Comunicación y Documentación, Universidad de Granada. Granada, España. benjamin@ugr.es

*** PhD Student. Statistical Cybermetrics Research Group, School of Computing and Information Technology, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1SB, United Kingdom. D.MinguilloBrehaut@wlv.ac.uk

**** Doctora en Documentación e Información Científica. Miembro del Grupo SCImago. Científica Titular del Consejo Superior de Investigaciones Científicas, Centro de Ciencias Humanas y Sociales, Instituto de Políticas y Bienes Públicos. 28037 Madrid, España. Zaida.Chinchilla@cchs.csic.es

***** Doctor en Filosofía y Letras. Investigador Principal del Grupo SCImago. Profesor de investigación del Consejo Superior de Investigaciones Científicas, Centro de Ciencias Humanas y Sociales, Instituto de Políticas y Bienes Públicos. 28037 Madrid, España. felix.demoya@cchs.csic.es

Cómo citar este artículo: VARGAS-QUESADA, Benjamín; MINGUILLO, David; CHINCHILLA-RODRÍGUEZ, Zaida; MOYA-ANEGÓN, Félix de. Estructura de la colaboración científica española en Biblioteconomía y Documentación (Scopus 1999-2007). *Revista Interamericana de Bibliotecología*. Ene.-Jun. 2010, vol. 33, no. 1, p. 105-123

Artículo recibido: 7 de abril de 2010. **Aprobado:** 24 de mayo de 2010

Abstract

Based on documents filed in the Scopus database between 1999 and 2007, this paper identifies, displays, and analyzes the main structures of scientific collaboration by authors in the library and information sciences in Spain. We apply techniques of similarity, techniques of social network analysis, like K-core, bibliometric procedures, such as author co-occurrence and an analysis of the frequency of terms in order to explore the communities of collaboration (subnets) that make up the most important components in this area. Through this we identify 21 components composed of 28 subnets that reveal the micro-structure of a network of scientific collaboration. The network showed collaboration that was centralized in a few components. This collaboration originated in the research sector and included information retrieval and information management. This paper highlights the limited cooperation between subnets in which institutional and geographical proximity play an important role and which have scientific production and evaluation as common fronts of collaboration. The methodology used manages to present an overview of the main collaborative scientific structure of authors in Spain during the time period under study.

Keywords: scientific collaboration, co authorship, collaboration networks, social network analysis, global centrality, library and information sciences, Spain

How to cite this article: VARGAS-QUESADA, Benjamín; MINGUILLO, David; CHINCHILLA-RODRÍGUEZ, Zaida; MOYA-ANEGÓN, Félix de. Structure of Spanish scientific collaboration in library and information sciences (Scopus 1999-2007). *Revista Interamericana de Bibliotecología*. Ene.-Jun. 2010, vol. 33, no. 1, p. 105-123

1. Introducción

La comunidad científica estructura sus relaciones conforme a modelos de redes sociales. En una red social, los actores (nodos) representan individuos, instituciones, disciplinas científicas, etc., y los enlaces, al conocimiento que intercambian dichos nodos. La principal aportación de las redes sociales estriba en que no se basan en un análisis individualista de las características de los actores, sino que se sustentan en la información relacional de los actores que componen la estructura científica (Vargas-Quesada y Moya-Anegón 2007).

La colaboración científica es un fenómeno social complejo de la investigación que se viene estudiando de forma sistemática desde finales de los 60. La coautoría

es una de sus formas más evidentes y mejor documentadas. Prácticamente cualquiera de sus aspectos puede ser estudiado de una forma fiable mediante el análisis de redes basado en métodos bibliométricos (Glänzel y Schubert, 2004). No obstante, los análisis de coautoría (micro) no son muy comunes (Newman, 2004), quizá debido a la complejidad en la recuperación y normalización de los datos. A nivel institucional (meso) y por países (macro) encontramos abundante literatura, pero estos estudios no reflejan de una forma adecuada los patrones y tendencias de colaboración entre individuos (Hou *et al.*, 2008; Kretschmer 2004). En este artículo combinamos indicadores bibliométricos con indicadores de red con los siguientes objetivos: 1) mostrar la micro-estructura de las redes de colaboración científica en el área de la Biblioteconomía y la Documentación en España, registrada en la base de datos Scopus (Elsevier, 2005); 2) identificar los principales grupos de investigación de dicho dominio; y 3) poner de manifiesto sus distintos frentes de colaboración.

2. Material y metodología

Puesto que nuestro objetivo es representar la estructura de la colaboración científica en Biblioteconomía y Documentación en España, la fuente de datos deberá ser lo suficientemente representativa y homogénea como para que la cobertura de datos que se van a analizar sea lo suficientemente razonable y estadísticamente fiable. El criterio de qué es investigación en Biblioteconomía y Documentación viene dado en parte por las propias materias que se tratan en el área, es decir: serán aquellos trabajos cuyo objeto de estudio pueda ser clasificado como una de las materias propias del dominio de la Biblioteconomía y Documentación (Järvelin and Vakkari, 1993). No obstante, no es fácil identificar las publicaciones científicas que cumplen este criterio. De hecho, éste es un problema típico de los estudios basados en publicaciones. Por ello, la identificación de las publicaciones científicas de un determinado campo o área suele estar basada en criterios subjetivos establecidos a priori por los propios autores: (Järvelin and Vakkari, 1993; Delgado López-Cozar, 2002; White and McCain, 1998); o en la adopción de clasificaciones de revistas hechas por grandes distribuidores: (Moya-Anegón de *et al.*, 2007; SCImago Grp. 2005, 2006).

Por esa razón, y con fines exclusivos de investigación, el 5 de junio de 2008 se descargaron de la base de datos Scopus todos los documentos publicados en las revistas, que tuviesen asignada la categoría Library and Information Science registradas en el SJR (SCImago 2007) y de cuyos autores uno, al menos, fuera español, es decir que en el campo "AFFIL" apareciese el término "spain" durante el periodo 1999-2007. La información obtenida para cada uno de los documentos ha sido:

autor(es), afiliación, título del documento, palabras clave, año de publicación, título de la revista fuente, volumen, número, páginas, número de citas recibidas y tipo de documento. Esta información ha sido exportada a una base de datos relacional.

No todos los autores firman sus trabajos de la misma forma a lo largo de su carrera investigadora. Por tanto, ha sido necesaria una labor de normalización de sus nombres. Para ello se ha empleado una herramienta interactiva que combina técnicas de minería de datos y visualización de redes, denominada D-Dupe (Bilgic *et al.*, 2006). La reducción ha sido significativa, y se ha pasado de 849 autores sin normalizar a 810 normalizados. Para la obtención de indicadores de red, identificación de grupos mediante k-core y el análisis de las redes de coautoría se ha utilizado Ucinet (Borgatti *et al.*, 2002). Para la visualización y representación gráfica de las redes se ha recurrido a Pajek (Batagelj and Mrvar, 2010). Finalmente, para la extracción de términos de los títulos de los documentos se ha utilizado un programa gratuito para el análisis de textos denominado Concordance (Watt, 2004).

En este trabajo representamos y analizamos la estructura de la colaboración científica española, recogida en la base de datos Scopus desde 1999 hasta 2007, ambos años incluidos. Para ello combinamos procedimientos de ARS con otros bibliométricos, como la coocurrencia de autores y el análisis de frecuencia de términos obtenidos de títulos y palabras clave.

Las redes de coautoría son del tipo *smallworld* y se caracterizan por un alto grado de *clustering* y una distancia geodésica media-baja en sus nodos (Newman, 2004). Este principio ha sido corroborado por Barabasi *et al.* (2002) y ampliado por Kretschmer y Aguillo (2004), y demuestra que en este tipo de redes existe, además, una relación inversa entre producción científica y distancia geodésica. Resulta por tanto evidente que los actores centrales, es decir aquellos con una distancia geodésica baja, tienen un papel clave en la transferencia e intercambio de información dentro de la red (Otte and Rousseau, 2002; Wasserman and Faust, 1998). De acuerdo con esto, en una red de colaboración, los nodos con un mayor grado tenderán a ocupar posiciones centrales y aquellos con los que más colaboran, se situarán cerca de ellos. A mayor número de colaboradores, mayor cercanía. Por tanto, si en una red de colaboración, por motivos de visualización y análisis, se eliminan de forma progresiva los autores con una menor cercanía, se estará prescindiendo de aquellos que ocupan posiciones periféricas, es decir, los que tienen una distancia geodésica elevada. En definitiva, se prescindirá de aquellos con un menor número de colaboraciones. Este procedimiento no es nuevo. Ya fue apuntado por Stokman y Snijders en un trabajo sin publicar (Scott, 2000), denominándolo “centralidad global”. Su objetivo es detectar un punto nuclear en la red y ordenar el resto de puntos a partir de su centralidad, para luego detectar una división “na-

tural” entre el centro, el margen y la periferia del gráfico. No obstante, implica el establecimiento de un umbral arbitrario de cercanía, que debe ser definido por el usuario o investigador, para indicar la división natural.

3. Resultados y discusión

3.1 Resultados básicos

De las 109 posibles revistas de la categoría LIS contenidas en Scopus, sólo 50 (46%) contenían al menos un trabajo firmado por algún autor afiliado a una institución española entre 1999 y 2007 (Ver **Anexo 1**). El número total de documentos descargados ha sido de 443, de los cuales 81 (18%) están firmados por un solo autor y 362 (82%) en colaboración. De estos últimos, 130 (36%) han sido firmados por dos autores, 122 (34%) por tres y 110 (30%) por más de tres. El número total de autores es de 810, de los cuales 50 (6%) son autores en solitario y 760 (94%) han colaborado con algún otro, dando lugar a un total de 1.231 pares de coautorías entre ellos. La media de documentos por autor es de 0,55 y la de coautoría por documento es de 2,10 (Ver **Tabla 1**).

Tabla 1. Datos básicos de la red de colaboración

| | | |
|---|------|-----|
| Número total de revistas | 109 | |
| Revistas con algún documento realizado por algún autor en España | 50 | 46% |
| Revistas sin ningún documento realizado por algún autor en España | 59 | 54% |
| Número total de documentos | 443 | |
| Documentos realizados por un único autor | 81 | 18% |
| Documentos realizados en co-autoría | 362 | 82% |
| Documentos realizados por dos autores | 130 | 36% |
| Documentos realizados por tres autores | 122 | 34% |
| Documentos realizados por más de tres autores | 110 | 30% |
| Número total de autores | 810 | |
| Autores sin ninguna colaboración | 50 | 6% |
| Autores con alguna colaboración | 760 | 94% |
| Autores con un único trabajo | 632 | 78% |
| Autores con dos trabajos | 98 | 12% |
| Autores con tres o más trabajos | 80 | 10% |
| Número total de colaboraciones | 1231 | |
| Media de documentos por autor | 0,55 | |
| Media de autores por documento | 2,10 | |
| Número total de autores con 3 o más trabajos en colaboración | 74 | 10% |

3.2 Estructura de la red: componentes principales

A partir de la información descargada y exportada a la base de datos relacional, se ha construido una matriz simétrica de 760 x 760 autores con, al menos, un trabajo en colaboración. Mediante Ucinet, se han calculado las medidas de Cercanía, Grado, Intermediación (normalizadas con valor máximo) y K-core de los 760 autores. Una vez calculadas estas medidas se inicia el procedimiento de identificación por centralidad global mediante un umbral de cercanía superior a 0,98891, teniendo en cuenta que los autores con tres o más colaboraciones, aunque su grado de cercanía esté por debajo del umbral establecido, no serán eliminados. Mediante este doble criterio se consigue representar la estructura de la colaboración científica de la red, identificando los autores con un alto nivel de colaboración y los componentes y subredes que estos mismos integran y lideran.

La estructura principal de la red de colaboración queda conformada por 21 componentes que integran a 215 coautores, que representan al 26,5% de los coautores de toda la red. Para su representación se ha utilizado el programa Pajek, recurriendo al algoritmo de *Kamada-Kawai* (1989) para la distribución de la información. No obstante, las posiciones de algunos autores se han modificado ligeramente para facilitar su visualización. El tamaño de los nodos de cada autor es proporcional al número de colaboradores del autor, mientras que el grosor de las líneas representa el número de trabajos en colaboración por cada par de autores (Ver **Figura 1**).

El componente 1 es el núcleo principal de la red. Cuenta con las colaboraciones más productivas e integra a 52 coautores. El componente 2 está formado por 40 coautores. El resto está constituido por un componente de 24 coautores, dos de 16, uno de 15, uno de 12, uno de 11 y uno de 10 coautores. Además existen 12 componentes periféricos que cuentan con entre 1 y 3 coautores (Ver **Tabla 2**).

3.3 Medidas de centralidad en la red de colaboración

El grado, en el caso de la colaboración científica, como sostienen Hou, Kretschmer y Liu (2008), es igual al número de colaboradores que un autor puede llegar a tener. Así, un autor con un mayor grado ha tenido que colaborar con un mayor número de autores. Esto implica que el autor es un actor central en la red. La cercanía es el promedio de la distancia geodésica entre un autor y el resto de actores de la red (Freeman, 1979). Un autor con un alto grado de cercanía podrá interactuar fácilmente con los actores relevantes de la red. Para ello, es necesario pertenecer a una subred de colaboración extensa y productiva. La intermediación refleja la medida en que un actor se beneficia de la posición de encontrarse en el camino más corto (de toda la red) que une dos actores (Freeman, 1977; Granovetter, 1973),

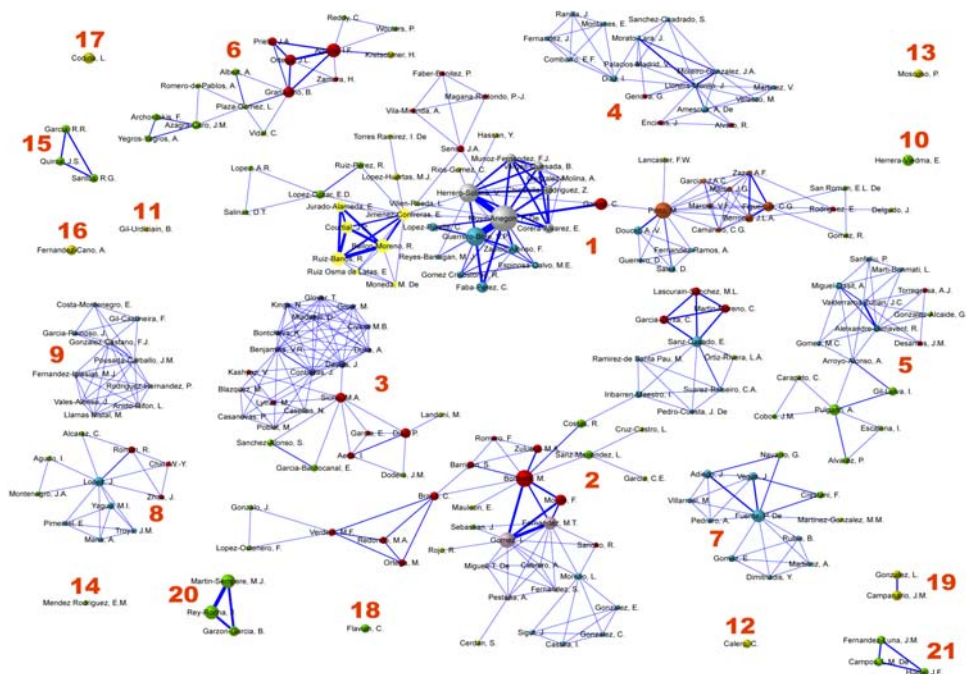


Figura 1. Red de colaboración científica española en Biblioteconomía y Documentación (Scopus 1998-2007). Disponible en tamaño real en: <http://www.scimago.es/benjamin/Figura-1.jpg>

Tabla 2. Componentes y número de coautores que los integran

| Comp. | Autor | Colaboraciones | N. de autores | % | Comp. | Autor | Colaboraciones | N. de autores | % |
|-------|------------------------|----------------|---------------|----|-------|------------------------|----------------|---------------|---|
| 1 | Moya-Anegón, F. de | 29 | 52 | 23 | 12 | Calero, C. | 3 | 2 | 1 |
| 2 | Gómez, I. | 10 | 40 | 17 | 13 | Moscoso, P. | 3 | 4 | 2 |
| 3 | Sicilia, M.A. | 3 | 24 | 10 | 14 | Méndez Rodríguez, E.M. | 3 | 5 | 2 |
| 4 | Llorens-Morillo, J. | 5 | 16 | 7 | 15 | García, R.R. | 3 | 3 | 1 |
| 5 | Pulgarín, A. | 5 | 16 | 7 | 16 | Fernández-Cano, A. | 4 | 2 | 1 |
| 6 | Aguillo, I.F. | 8 | 15 | 7 | 17 | Codina, L. | 4 | 3 | 1 |
| 7 | Fuente, P. De | 7 | 12 | 5 | 18 | Flavián, C. | 3 | 3 | 1 |
| 8 | López, J. | 4 | 11 | 5 | 19 | Campanario, J.M. | 4 | 2 | 1 |
| 9 | González-Castaño, F.J. | 2 | 10 | 4 | 20 | Martín-Sempere, M.J. | 8 | 3 | 1 |
| 10 | Herrera-Viedma, E. | 4 | 2 | 1 | 21 | Campos, L.M. de | 3 | 3 | 1 |
| 11 | Gil-Urdiciain, B. | 3 | 2 | 1 | | | | | |

pudiendo controlar el flujo de información y por tanto tener un papel importante en la intermediación de autores que de otra forma no tendrían enlaces directos entre sí (Zhu *et al.*, 2008). Si observamos la **tabla 3** veremos que los autores más centrales, que más fácilmente interactúan con el resto y que más controlan el flujo de información en la red son: Moya-Anegón, Pinto, Jiménez-Contreras, Gómez, Bordóns y Sanz-Casado.

3.4 Comunidades de colaboración: subredes

Una vez identificada la estructura de la red a partir de sus componentes, se pasa a identificar, mediante K-core, las subredes que conforman cada uno de esos componentes, identificando así las distintas comunidades de colaboración de la red. En la **figura 1** se muestran en distinto color, para que puedan ser fácilmente identificadas. K-core es una medida basada en el nivel de cohesión que se produce entre los actores a partir de su grado, buscando el mayor número de autores altamente conectados entre sí (Ortega y Aguillo, 2007; Scott, 2000). Mediante este procedimiento, se identifican hasta 28 subredes distintas y 38 autores principales (Ver **Tabla 3**). Cada subred está identificada a partir del autor con mayor número de trabajos. No obstante, aparecen seis subredes con más de un autor central (marcados con *), debido a que sus autores tienen el mismo grado y número de colaboraciones.

Los componentes 1 y 2 son los que mayor número de subredes integran: 3. Encontramos también 3 componentes con dos subredes. El resto de componentes, 16, son una única subred. La cooperación entre las distintas comunidades de colaboración es realmente escasa, ya que solo 12 de las 28 subredes han logrado cooperar con otra subred. Esto apunta a que sólo 5 de los 21 componentes son lo suficientemente amplios y sólidos como para unir más de dos comunidades de investigadores distintas, aunque también deja abierta la hipótesis de que la mayoría de los componentes prefieren investigar en frentes distintos y no colaborar con el resto.

3.5 Correlación entre producción y medidas de centralidad

A partir del cálculo del índice de correlación de *Pearson* entre la producción en colaboración de los 760 autores y las tres principales medidas de centralidad: $r = 0.515$, 0.640 y $-0,403$ (Ver **Tabla 4**), se obtiene una correlación positiva entre el grado y la intermediación con la producción de los autores, lo que demuestra una clara relación entre los autores productivos y su capacidad para colaborar con el resto de autores. Mientras que la correlación débil y negativa entre cercanía y producción, confirma la relación inversa de estas medidas apuntada por Kretschmer (2004): a mayor producción menor distancia con el resto de actores.

Tabla 3. Principales autores de las subredes

| Comp. | Subred | Autor | Grado Norm. | Intermediación Norm. | Cercanía Norm. | |
|-------|--------|-------------------------|-------------|----------------------|----------------|---|
| 1 | 1.1 | Moya-Anegón, F. de | 1 | 1 | 0,93720 | |
| | 1.2 | Pinto, M. | 0,7647 | 0,97943 | 0,93725 | |
| | 1.3 | Jiménez-Contreras, E. | 0,6471 | 0,81224 | 0,93723 | |
| 2 | 2.1 | Gómez, I. | 0,6471 | 0,32991 | 0,95199 | |
| | 2.2 | Bordóns, M. | 0,5882 | 0,84968 | 0,95198 | |
| | 2.3 | Sanz-Casado, E. | 0,4706 | 0,19778 | 0,95208 | |
| 3 | 3.1 | Contreras, J. | 0,7647 | 0,05696 | 0,97166 | |
| | | Benjamins, V.R. | 0,7647 | 0,05696 | 0,97166 | * |
| | 3.2 | Sicilia, M.A. | 0,4706 | 0,19304 | 0,97166 | |
| 4 | 4.1 | Llorens-Morillo, J. | 0,6471 | 0,08333 | 0,98151 | |
| 5 | 5.1 | Alexandre-Benavent, R. | 0,5294 | 0,03613 | 0,98151 | |
| | | Valderrama-Zurián, J.C. | 0,5294 | 0,03613 | 0,98151 | * |
| | 5.2 | Pulgarín, A. | 0,3529 | 0,07595 | 0,98152 | |
| 6 | 6.1 | Aguillo, I.F. | 0,4118 | 0,05591 | 0,98275 | |
| | 6.2 | Plaza-Gómez, L. | 0,2941 | 0,06408 | 0,98275 | |
| 7 | 7.1 | Fuente, P. de | 0,6471 | 0,06013 | 0,98643 | |
| 8 | 8.1 | López, J. | 0,5882 | 0,05222 | 0,98766 | |
| 9 | 9.1 | González-Castaño, F.J. | 0,5294 | 0,01187 | 0,98890 | |
| | | Pousada-Carballo, J.M. | 0,5294 | 0,01187 | 0,98890 | * |
| 10 | 10.1 | Herrera-Viedma, E. | 0,3529 | 0,02532 | 0,99137 | |
| 11 | 11.1 | Gil-Urdiciain, B. | 0,1176 | 0,00949 | 0,99138 | |
| 12 | 12.1 | Calero, C. | 0,3529 | 0,01503 | 0,99260 | |
| 13 | 13.1 | Moscoso, P. | 0,1765 | 0,01424 | 0,99260 | |
| 14 | 14.1 | Méndez Rodríguez, E.M. | 0,2941 | 0,01266 | 0,99383 | |
| 15 | 15.1 | García, R.R. | 0,2941 | 0,00106 | 0,99383 | |
| | | Quirós, J.S. | 0,2941 | 0,00106 | 0,99383 | * |
| | | Santos, R.G. | 0,2941 | 0,00106 | 0,99383 | * |
| 16 | 16.1 | Fernández-Cano, A. | 0,2353 | 0,00791 | 0,99506 | |
| 17 | 17.1 | Codina, L. | 0,2353 | 0,00554 | 0,99506 | |
| 18 | 18.1 | Flavián, C. | 0,1765 | 0,00316 | 0,99630 | |
| 19 | 19.1 | Campanario, J.M. | 0,1765 | 0,00316 | 0,99630 | |
| 20 | 20.1 | Martín-Sempere, M.J. | 0,1176 | 0,00000 | 0,99753 | |
| | | Rey-Rocha, J. | 0,1176 | 0,00000 | 0,99753 | * |
| 21 | 21.1 | Campos, L.M. de | 0,1176 | 0,00000 | 0,99753 | |
| | | Fernández-Luna, J.M. | 0,1176 | 0,00000 | 0,99753 | * |
| | | Huete, J.F. | 0,1176 | 0,00000 | 0,99753 | * |

Tabla 4. Correlación entre la producción y las principales medidas de centralidad

| | Producción | Grado | Intermediación | Cercanía |
|----------------|------------|--------|----------------|----------|
| Producción | 1 | | | |
| Grado | 0,515 | 1 | | |
| Intermediación | 0,640 | 0,408 | 1 | |
| Cercanía | -0,403 | -0,412 | -0,349 | 1 |

3.6 Frentes de colaboración de la red

Con la ayuda del software Concordance se ha analizado la frecuencia de los términos extraídos de los títulos y palabras clave de los 362 documentos publicados en colaboración. De esta forma es posible identificar los principales frentes de investigación en Biblioteconomía y Documentación en España a nivel internacional (Ver **Tabla 5**). Destaca la “recuperación” y “gestión de la información”, así como la propia “documentación” y los “sistemas de información”. También se observa un fuerte interés en torno a la “comunicación científica” y otros frentes como: “Internet”, “e-learning” y “la biblioteca y el entorno digital y electrónico”. El escaso interés que por sí sola despierta la biblioteca confirma una producción asimétrica entre los sectores profesional e investigador, acentuada por la limitada colaboración entre ellos (Cano, 1999; Delgado López-Cózar, 2002). En cuanto al nivel territorial destacan los estudios del ámbito nacional.

Tabla 5. Términos más frecuentes en toda la red de colaboración

| Términos | Frecuencia | Términos | Frecuencia |
|--------------------------|------------|----------------------|------------|
| Information | 233 | Model (s) | 62 |
| System (s) (-atic) | 177 | Knowledge | 57 |
| Science (s) (-tific) | 130 | Learning | 53 |
| Web | 95 | Research | 50 |
| Analyse (s) (-ing) | 93 | Library (-ies) | 50 |
| Retrieval | 85 | Electronic | 48 |
| Spain (-nish) | 82 | Study (-ies) | 47 |
| Management (-ers) (-ing) | 71 | Process (-es) (-ing) | 43 |
| Computer | 66 | Digital | 40 |

3.7 Frentes de colaboración de las distintas subredes

Al igual que en el caso anterior, también se ha analizado la frecuencia de los términos extraídos de los títulos y palabras clave de los trabajos publicados por los autores de cada una de las 28 subredes detectadas. Así, una vez identificada la institución a la que pertenecen los autores de cada subred, sus frentes de colaboración y la intensidad con la que colaboran, se pueden determinar tres tipos de redes o componentes.

1. Red institucional: es aquel componente formado por subredes de actores que pertenecen básicamente a una institución común y en la que cada subred cuenta con frentes de investigación diferenciados. Entre la subredes se colabora de forma esporádica mediante un número reducido de autores y no en base a un frente de colaboración común sino, probablemente, por motivos de proximidad institucional.

2. Red temática: es aquel componente formado por subredes de actores que pertenecen básicamente a múltiples instituciones y en la que cada subred cuenta con frentes de investigación similares o comunes. Estas subredes colaboran mediante un mayor número de autores; son redes más cohesionadas, que colaboran evidentemente en torno a un frente temático de colaboración común, independientemente de la cercanía física e institución a la que pertenezcan los autores.
3. Red territorial: es aquel componente formado por subredes de actores que pertenecen a varias instituciones y que se caracterizan por estar cercanas geográficamente, lo que da lugar a una colaboración temática un tanto heterogénea, establecida probablemente por la cercanía física de los investigadores.

Volviendo a la **Figura 1** y a la **Tabla 3**, podemos decir que los componentes principales de la red son el 1 y el 2, los cuales mostramos de forma separada para su mejor análisis. El componente 1 (Ver **Figura 2**), está formado por tres subredes estructuradas en torno a tres autores de la Universidad de Granada (UGR). La subred principal (1.1) de este componente, está liderada por Moya-Aneón y la conforma básicamente el grupo de investigación SCImago.

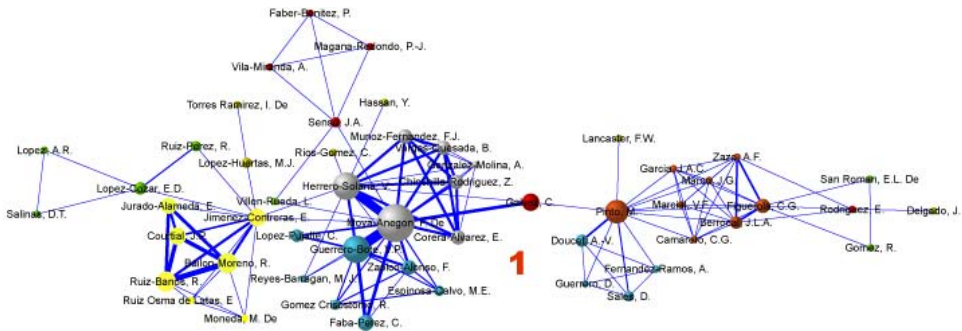


Figura 2. Componente 1. Disponible en tamaño real en: <http://www.scimago.es/benjamin/Figura-2.jpg>

Sus frentes de colaboración más relevantes son: la recuperación de información mediante técnicas de visualización de información, el análisis de dominios científicos y la representación mediante “self-organizing maps”. Dentro de esta subred se diferencian dos grupos altamente cohesionados, con autores pertenecientes a la UGR y a la Universidad de Extremadura (UNEX). Este último grupo está liderado por Guerrero-Bote y sus colaboraciones se concentran en el desarrollo de algoritmos genéticos en un entorno web y en los “self-organizing maps”. Esta subred (1.1) se une a la liderada por Pinto (1.2) debido a una colaboración entre ésta y

Gálvez. Los frentes de colaboración de esta subred son un tanto heterogéneos y entre ellos se destacan la evaluación de recursos en torno a la educación superior en España y la recuperación de información. En la subred (1.2) también se distinguen dos grupos de autores unidos por Pinto, uno, procedente de la UGR y otro, básicamente de la Universidad de Salamanca (Grupo Reina). La subred (1.3) está liderada por Jiménez-Contreras y se une a la subred (1.1) por una colaboración de este con Moya-Anegón y Herrero-Solana. Se trata de una subred de colaboración compuesta por miembros del grupo de investigación Ec3 de la UGR, especializada en la evaluación de revistas científicas. En esta misma subred es posible identificar un subgrupo que trabaja en torno a la catalogación y el fraude científico. Destacar que la colaboración entre las tres subredes es prácticamente inexistente a pesar de estar conformadas, en su mayoría, por investigadores de la propia UGR, que puede considerarse una red claramente institucional. Los temas principales de colaboración de la red son la recuperación y visualización de información, y el análisis bibliométrico a partir de la cocitación.

El componente 2 (Ver **Figura 3**) está formado por tres subredes lideradas por dos autoras del CINDOC (actual IEDCYT) y un autor de la Universidad Carlos III, de Madrid (UC3M).

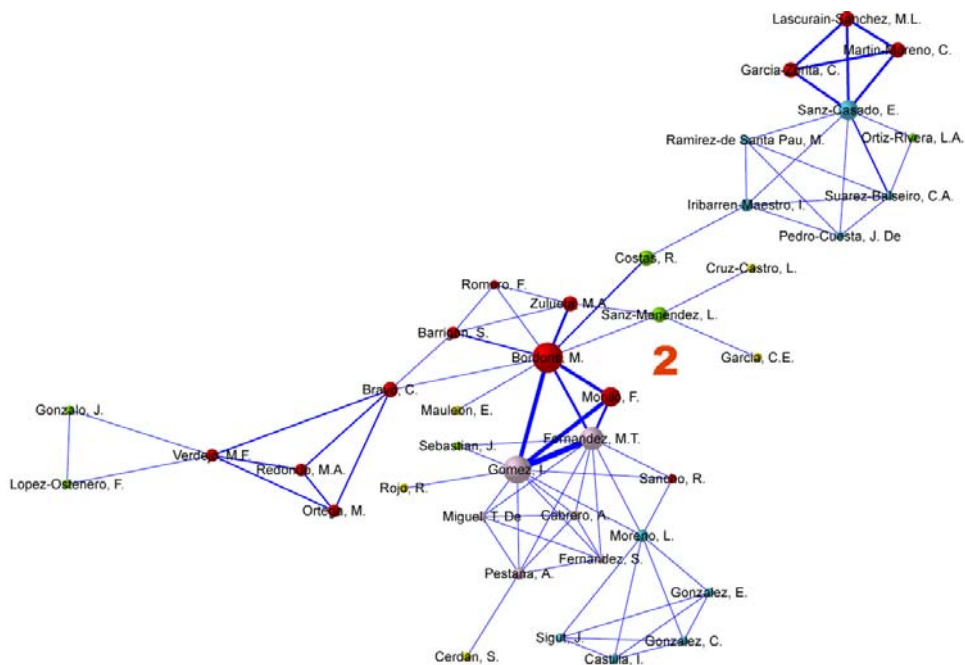


Figura 3. Componente 2. Disponible en tamaño real en: <http://www.scimago.es/benjamin/Figura-3.jpg>

En la subred principal (2.1) la autora más representativa en términos de colaboración es Isabel Gómez. Esta subred está compuesta básicamente por un grupo de investigadores del CINDOC que tienen como frentes de colaboración el análisis bibliométrico de la producción científica española y el desarrollo de indicadores bibliométricos. Esta subred colabora con un pequeño grupo procedente de la Universidad Complutense de Madrid. La segunda subred (2.2) se relaciona con la subred (2.1), debido a la extraordinaria colaboración en estudios bibliométricos que existe entre Isabel Gómez y Fernández-Muñoz con Morillo y Bordóns. Esta última autora es quien encabeza esta subred, formada por autores del CINDOC y por miembros de algunas universidades de la Comunidad de Madrid. Sus frentes de colaboración son la recuperación de información y las herramientas bibliométricas. La tercera subred (2.3) la componen autores de la UC3M liderados por Sanz-Casado. Esta subred se une a la 2.2 mediante una colaboración entre Iribarren-Maestro con Costas y dos colaboraciones de Costas con Bordóns sobre indicadores bibliométricos. Los frentes de colaboración de esta subred son un tanto heterogéneos. Podrían destacarse los estudios sobre producción científica y los mapas bibliométricos. El componente 2 es una red institucional y territorial que se caracteriza por estar integrada básicamente por investigadores del CINDOC que colaboran con autores pertenecientes a universidades de la Comunidad de Madrid. Entre sus frentes de colaboración destacan los análisis bibliométricos.

El componente 3 (volviendo a la **Figura 1** de ahora en adelante), se crea a partir de dos subredes lideradas por autores del sector privado y de la Universidad de Alcalá de Henares (UAH). La subred principal (3.1) de este componente, está liderada por Jesús Contreras y Benjamins, pertenecientes a la empresa Intelligent Software Components, de Madrid, y se compone de autores del sector privado, la Universidad Autónoma de Barcelona, y algunas universidades extranjeras. Esta subred, a pesar de ser extensa, tan sólo se basa en dos trabajos; se puede destacar como tema de colaboración en ella la gestión del conocimiento. Las subredes 3.1 y 3.2 se unen mediante un trabajo en colaboración con Sicilia, de la UAH. Justamente este autor lidera la subred 3.2, compuesta en su mayoría por autores de la UC3M, la UAH y universidades extranjeras que colaboran en torno a la gestión de conocimiento y la web semántica. La colaboración observada entre estas dos subredes se caracteriza por integrar autores procedentes de múltiples instituciones que colaboran en torno a un determinado campo de investigación, diferenciándose así del resto de subredes analizadas, entre las que destaca una mayor colaboración, debida posiblemente a la proximidad institucional y territorial, y que conforman, por tanto, una red o componente temático.

El componente 4 está formado por una sola subred (4.1), con Llorens-Morillo, de la UC3M, a la cabeza. Los autores que la conforman pertenecen esencialmente

al área de informática y provienen básicamente de la UC3M. Sus frentes de colaboración son la recuperación de información, la clasificación e indexación automatizada, ontologías y tesauros.

El componente 5 se constituye a partir de dos subredes lideradas por un autor del Instituto de Historia de la Ciencia y Documentación (IHCD) y otro de la UNEX. La subred 5.1 está encabezada por Aleixandre-Benavent y la conforman básicamente autores del IHCD y de Hospitales Universitarios de Valencia. Sus trabajos en colaboración tratan sobre el análisis bibliométrico de revistas médicas en España. La subred 5.1 se une a la 5.2 mediante la colaboración entre Arroyo-Alonso, de la Universidad de Valencia, con Gil-Leiva, de la Universidad de Murcia, y Pulgarín de la UNEX. Justamente este último autor es el más visible de la subred 5.2, centrada en el análisis bibliométrico. Este componente es una red temática que integra autores procedentes de Valencia y Extremadura en torno a la bibliometría.

El componente 6 está constituido por dos subredes lideradas por autores del CINDOC. La subred 6.1 está integrada por autores pertenecientes al grupo de investigación de Cibermetría del CINDOC y algunos autores afiliados a instituciones holandesas. Esta subred está dirigida por Aguillo, y sus principales frentes de colaboración son la webmetría y el análisis de redes sociales. La subred 6.2 está canalizada por Plaza-Gómez y la componen autores del CINDOC, del Instituto de Gestión de la Innovación y del Conocimiento (INGENIO) adscrito al CINDOC, y a la Universidad Politécnica de Valencia. Estos autores colaboran en el análisis de la producción científica y de patentes a nivel nacional. Ambas subredes se unen mediante dos colaboraciones: una entre Granadino con Plaza Gómez y Vidal, del CINDOC, y otra entre Granadino con Plaza Gómez y Albert, que también pertenecen a la misma institución. El componente 6 es una red institucional, integrada fundamentalmente por investigadores del CINDOC que colaboran en torno a estudios de webometría y bibliometría.

En el componente 7 se encuentra Pablo de la Fuente, de la Universidad de Valladolid (UVA), y sus autores pertenecen al Departamento de Informática de la UVA. Sus principales frentes de colaboración son la compresión y recuperación de documentos estructurados, indexación de información e interfaces de usuario. El componente 8 se establece a partir de una subred de autores del Grupo de Ingeniería de Software de la Universidad de Málaga (GISUM) y dos autores de Singapur. La encabeza Javier López y su principal tema de colaboración es la seguridad de la información.

El componente 9 es una subred integrada básicamente por autores del Grupo de Tecnologías de la Información (GTI), del Departamento de Ingeniería Telemática de la Universidad de Vigo. Esta subred se origina en dos trabajos en colaboración

que arrojan como autor más representativo a González-Castaño, y sus trabajos tratan sobre la aplicación de las nuevas tecnologías en la educación.

Los siguientes componentes representan al autor o autores que tienen tres o más trabajos en colaboración. Se trata de redes de colaboraciones reducidas y periféricas. Al componente 10 le da origen Herrera-Viedma, del Departamento de Ciencias de la Computación de la UGR. En este componente colaboran básicamente autores afiliados a los Departamentos de Ciencias de la Computación y Biblioteconomía y Documentación de la UGR. Tiene como principal frente de colaboración la utilización de métodos difusos para la evaluación y recuperación de información. El componente 11 está liderado por Gil-Urdiciain, de la Universidad Complutense de Madrid (UCM), y está compuesto por autores de la (UCM) y del CINDOC. Sus frentes de colaboración son diversos y destacan la evaluación cualitativa de resúmenes, bases de datos y blogs. El componente 12 fue iniciado por Calero y lo integran básicamente autores de la Universidad de Castilla la Mancha. Su frente de colaboración es el modelo de calidad web. El componente 13 está liderado por Moscoso,, de la UAH y lo integran autores de la UAH y de la UC3M. Su colaboración gira en torno a los estudios de los OPACs en España. El líder del componente 14 es Méndez Rodríguez, de la UC3M, al igual que los autores que integran el componente, y su colaboración gira en torno a la web semántica. El componente 15 está liderado por tres autores del Departamento de Construcción e Ingeniería de la Universidad de Oviedo que tienen como frente de colaboración la enseñanza del diseño asistido por ordenador (CAD). El componente 16 lo inicia Fernández-Cano y lo integran autores de la Facultad de Educación de la UGR. Su colaboración se centra en estudios cuantitativos sobre la investigación española en educación. El componente 17 fue creado por Codina y autores del grupo de investigación DigiDoc de la Universidad Pompeu Fabra, que colaboran en torno a la web semántica y software de acceso abierto. El componente 18 está compuesto por Flavián y autores de la Facultad de Ciencias Económicas y Empresariales de la Universidad de Zaragoza, que colaboran en la seguridad y protección de datos electrónicos. El componente 19 lo conforma un pequeño grupo de autores del Departamento de Física de la UAH, destacando Juan Campanario y Lidia González. Su frente de colaboración se encuentra en el estudio del índice de impacto de las revistas científicas. El componente 20 es una triada que tuvo su origen en una alta colaboración entre Martín-Sempere, Rey-Rocha y Garzón-García, del CINDOC, que tienen como frente de colaboración el análisis en la producción de la investigación nacional en Geología. Finalmente, el componente 21 también es una triada que integra a Luis Campos, Fernández-Luna y Huete, del grupo de investigación de Tratamiento de la Incertidumbre en Inteligencia Artificial (UTAI) de la UGR, y que tienen como frente de colaboración las redes bayesianas y la recuperación de información.

A raíz del análisis pormenorizado de las subredes, se observa que la producción de la mayoría de las comunidades de investigadores es baja. De las 28 subredes 18 se han creado a partir de entre 3 y 6 trabajos en colaboración, en un periodo de nueve años. En cuanto a los frentes que atraen una mayor colaboración, se centran en la producción y evaluación científica, seguidos de la recuperación de la información, la web semántica, y la enseñanza. Sin embargo, debido a la heterogeneidad en los frentes de colaboración producidos por algunas subredes, se puede observar que la producción de ciertas comunidades de investigadores carece de frentes de colaboración o líneas de investigación claras y consolidadas en el tiempo. Un análisis dinámico permitiría establecer si existe correlación entre la consolidación de los frentes de investigación y el tiempo productivo de los autores.

Entre las instituciones que dan origen a un mayor número de subredes destacan la UGR, el CINDOC, la UC3M y la UHA. También se observan una serie de subredes iniciadas básicamente por instituciones relacionadas directamente con la informática. La colaboración entre las distintas subredes da lugar a componentes de tipo institucional y territorial.

4. Conclusiones

El aporte principal de este trabajo radica en la presentación de una metodología para la representación y análisis de la estructura de la colaboración científica, no sólo aplicable al área de la Biblioteconomía y Documentación, sino a cualquier área o dominio científico.

Mediante la utilización de medidas de similaridad, técnicas de análisis de redes sociales y análisis de frecuencia de términos, se ha identificado, representado y analizado la estructura principal de la red de colaboración científica de los autores de Biblioteconomía y Documentación recogidos en la base de datos Scopus en el periodo 1999-2007.

De los 760 coautores, tan sólo 74 (10%) tienen tres o más trabajos en colaboración. Mediante la centralidad global se identifican 215 coautores repartidos en 21 componentes que conforman la estructura principal de la red. Los componentes principales y consolidados son los 3 primeros, que logran integrar el 50% de autores representados. En cuanto a las comunidades de colaboración, se identifican 28 subredes, entre las que es notoria la reducida cooperación, ya que sólo 5 de los 21 componentes son capaces de unir dos subredes de investigadores distintas. Mientras que la correlación positiva indica que los autores más productivos tienen una mayor capacidad de colaboración, la correlación negativa indica la relación inversamente proporcional entre la producción y la distancia con el resto de los actores.

Entre los frentes de colaboración más importantes destacan: los estudios a nivel nacional sobre recuperación y gestión de información, la documentación, la evaluación de la producción científica, los sistemas de información, la enseñanza electrónica, y la biblioteca, que ocupa unos de los últimos puestos. Se puede afirmar que la producción internacional está procedente del sector investigador académico. Con respecto a las subredes, el frente que suscita una mayor colaboración es la producción y evaluación de la comunicación científica, seguida de la recuperación de información. Las instituciones que dan lugar a un mayor número de subredes son la UGR, el CINDOC, la UC3M y la UHA. La unión de estas subredes crea principalmente componentes de tipo institucional y territorial, donde se observa que un factor importante en la colaboración en la Biblioteconomía y Documentación española es la proximidad geográfica e institucional entre los colaboradores. El componente con mayor grado de colaboración de toda la red es el 1. originado en tres subredes de la UGR. Integra al 23% de los autores representados, entre los que destacan el autor principal de la red y cinco de las siete parejas más productivas. El componente 2 es el segundo en importancia en toda la red. Se constituye a partir de la colaboración de dos subredes del CINDOC y una de la UC3M. Integra el 17% de los autores, entre los que se encuentra una de las parejas más productivas de la red. Por todo ello, estos dos componentes se pueden considerar como las dos redes más relevantes e influyentes.

Los resultados que aquí aparecen no son más que una imagen estática de la colaboración científica a nivel micro. Sin embargo, la colaboración científica cambia a lo largo del tiempo, y eso no lo reflejan este tipo de investigaciones. Por ello, nuestro próximo objetivo es representar la evolución de la estructura de la colaboración científica en la misma área y base de datos a lo largo del tiempo.

Referencias bibliográficas

- BARABASI, A.L., et al. 2002. Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, vol. 311, no. 3-4, p. 590-614.
- BATAGELJ, V., and MRVAR, A. 2010. Pajek 1.26: *package for large network analysis* [en línea] [citado marzo 12, 2010]. Disponible en: <http://pajek.imfm.si>
- BILGIC, M. et al. 2006. *D-Dupe 2.0: a novel tool for interactive data deduplication and integration*. [en línea]. citado marzo 12, 2010] Disponible en: <http://www.cs.umd.edu/projects/linqs/ddupe/index.htm>.
- BORGATTI, S.P., EVERETT, M.G., and FREEMAN, L.C. 2002. *UCINET 6 for Windows: software for social networks analysis*. Harvard: Analytic Technologies.
- CANO, V. 1999. Bibliometric overview of library and information science research in Spain. *Journal of the American Society for Information Science*, vol. 50, no. 8, p. 675-680.
- DELGADO LÓPEZ-CÓZAR, E. 2002. *La investigación en Biblioteconomía y Documentación*. Gijón: Trea
- ELSEVIER 2005. *Scopus*. [en línea] Amsterdam: Elsevier, 2005 [citado marzo 12, 2010] Disponible en: <http://www.scopus.com>

- FREEMAN, L.C. 1977. A set of measures of centrality based on betweenness. *Sociometry*, vol. 40, no. 1, p. 35-41.
- FREEMAN, L.C. 1979. Centrality in social networks: conceptual clarification. *Social Networks*, vol. 1, no. 3, p. 215-239.
- GLÄNZEL, W., and SCHUBERT, A. 2004. Analyzing scientific networks through co-authorship. In: MOED, H., GLÄNZEL, W. and SCHMOCH, U., eds. *Handbook of quantitative science and technology research*. The Netherlands: Kluwer Academic Publishers.
- GRANOVETTER, M.S. 1973. The strength of weak ties. *American Journal of Sociology*, vol. 78, no. 6, p. 1360-1380.
- HOU, H.; KRETSCHMER, H., and LIU, Z. 2008. The structure of scientific collaboration networks in Scientometrics. *Scientometrics*, vol. 75, no. 2, p. 189-202.
- JÄRVELIN, K., and VAKKARI, P. 1993. The evolution of library and information science 1965-1985: a content analysis of journal articles. *Information Processing and Management*, vol. 29, no. 1, p. 129-144.
- KAMADA, T. and KAWAI, S. 1989. An algorithm for drawing general undirected graphs. *Information Processing Letters*, vol. 31, no. 1, p. 7-15.
- KRETSCHMER, H. 2004. Author productivity and geodesic distance in bibliographic co-authorship networks, and visibility on the Web. *Scientometrics*, vol. 60, no. 3, p. 409-420.
- KRETSCHMER, H., and AGUILLO, I. F. 2004. Visibility of collaboration on the Web. *Scientometrics*, vol. 61, no. 3, p. 405-426.
- MOYA-ANEGON, F. de, *et al.* 2007. Coverage analysis of Scopus: a journal metric approach. *Scientometrics*, vol. 73, no. 1, p. 53-78.
- NEWMAN, M. 2004. Who is the best connected scientist?: a study of scientific coauthorship networks. In: Ben-Naim, E.; Frauenfelder, H. and Toroczkai Z. (eds.). *Complex networks. Part III Information Networks & Social Networks*. Berlin: Springer, vol. 650, p. 337-370
- ORTEGA, L.L., y AGUILLO, I. F. 2007. La web académica española en el contexto del espacio europeo de educación superior: estudio exploratorio. *El Profesional de la Información*, vol. 16, no. 5, p. 417-425.
- OTTE, E. and ROUSSEAU, R. 2002. Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, vol. 28, no. 6, p. 441-453.
- SCImago RESEARCH GROUP. 2005. Producción española con visibilidad internacional (ISI-WOS) en Biblioteconomía y Documentación (I). *El Profesional de la Información*, vol. 14, no. 6, p. 459-461.
- SCImago RESEARCH GROUP. 2006. Producción española con visibilidad internacional (ISI-WOS) en Biblioteconomía y Documentación (II). *El Profesional de la Información*, vol. 15, no. 1, p. 34-36.
- SCImago RESEARCH GROUP. 2007. *SJR - SCImago Journal y Country Rank*. [en línea]. Madrid: SCImago, 2007. [citado marzo 12, 2010] Disponible en: <http://www.scimagojr.com>
- SCOTT, J. 2000. *Social network analysis: a handbook*. Los Angeles: Sage.
- VARGAS-QUESADA, B. and MOYA-ANEGÓN, F. 2007. *Visualizing the structure of science*. Berlin: Springer.
- WASSERMAN, S. and FAUST, K. 1998. *Social network analysis: methods and applications*. Cambridge: Cambridge University Press.
- WATT, R. J.C. 2004. *Concordance, text analysis and concordance software*.
- WHITE, H.D. and McCAIN, K.W. 1998. Visualizing a discipline: an author co-citation analysis of information science, 1972-1995. In: *Journal of the American Society for Information Science*, vol. 49, no. 4, p. 327-355.
- ZHU, W., CHEN, C. and ALLEN, B. 2008. *Analyzing the propagation of influence and concept evolution in enterprise social networks through centrality and latent semantic analysis*. Berlin: Springer.

Anexo I

Revistas de Biblioteconomía y Documentación de las que se ha descargado al menos un trabajo.

| Revistas | Nro. de docs | Revistas | Nro. de docs |
|---|--------------|--|--------------|
| Aslib Proceedings | 1 | Journal of Information Science | 15 |
| Cataloging and Classification Quarterly | 1 | Journal of Information Technology | 6 |
| Collection Building | 1 | Journal of Knowledge Management | 14 |
| Computers and Education | 19 | Journal of Librarianship and Information Science | |
| Development and Learning in Organisations | 1 | Journal of Management Information Systems | 3 |
| D-Lib Magazine | 7 | Journal of the American Society for Information Science and Technology | 1 |
| Electronic Library | 8 | Journal of the Medical Library Association | 4 |
| Government Information Quarterly | 5 | Knowledge Organization | 4 |
| Information Processing and Management | 28 | Lecture Notes in Control and Information Sciences | 4 |
| Information Research | 14 | Library and Information Science Research | 4 |
| Information Security Technical Report | 3 | Library Management | 10 |
| Information Society | 1 | Library Trends | 2 |
| Information Systems Journal | 1 | Libri | 4 |
| Information Technology and Libraries | 5 | MIS Quarterly: Management Information Systems | 8 |
| Information-Wissenschaft und Praxis | 1 | Online Information Review | 1 |
| Interlending and Document Supply | 2 | Profesional de la Informacion | 20 |
| International Information and Library Review | 1 | Research Evaluation | 60 |
| International Journal of Geographical | | Restaurator | 18 |
| Information Science | 7 | Scientist | 9 |
| International Journal of Information Management | 8 | Scientometrics | 7 |
| International Journal of Lexicography | 4 | Social Science Computer Review | 75 |
| Journal of Academic Librarianship | 3 | Social Science Information | 1 |
| Journal of Archival Organization | 1 | Telecommunications Policy | 2 |
| Journal of Digital Information | 1 | VINE | 5 |
| Journal of Documentation | 8 | | 1 |
| Journal of Enterprise Information Management | 3 | | |
| Journal of Health Communication | 2 | | |