

# Mapas de ciencias multidisciplinares: la biología molecular en la Comunidad de Madrid

Carlos Olmeda Gómez<sup>1</sup>, Antonio Perianes-Rodríguez<sup>1</sup>, M<sup>a</sup> Antonia Ovalle-Perandones<sup>1</sup>

<sup>1</sup>Grupo de investigación SCImago. Departamento de Biblioteconomía y Documentación. Universidad Carlos III de Madrid. C/ Madrid 128. 28903 Madrid. España. carlos.olmeda@uc3m.es

## Resumen

*En esta comunicación se presentan los resultados de un estudio bibliométrico sobre las relaciones estructurales en la clase temática de la Biología Molecular, Celular y Genética a partir de la producción científica en este campo en la Comunidad de Madrid, en el período comprendido entre 1995 y 2003, mediante un análisis de cocitación y de colaboración científica institucional, empleando técnicas de redes sociales. El objetivo del trabajo, ha sido examinar las relaciones disciplinares e institucionales que se producen en este campo científico. Para ello, se han construido mapas heliocéntricos de cocitación de categorías, mediante técnicas que descubren la estructura de la red intelectual de la disciplina y mapas de colaboración institucional para explorar la red social de instituciones que colaboran en la producción científica en este dominio.*

**Palabras clave:** Análisis de cocitación, Colaboración científica, Mapas de ciencia, Multidisciplinariedad.

## Abstract

*A cocitation and a coauthor analysis about Madrid's scientific production in Molecular and Cellular Biology and Genetics was conducted over a time period spanning the years 1995-2003. We use ISI-JCR categories as units of cocitation and measurement for the construction of heliocentric maps. Applying visualization techniques and social network analysis, we presented three maps. This analysis reveals the changing linkages of the domain and the multidisciplinary intellectual structure of the field.*

**Keywords:** Cocitation analysis, Maps of Science, Multidisciplinarity, Scientific collaboration.

## 1 Introducción

Los métodos cuantitativos que se usan para el estudio de la interdisciplinariedad y multidisciplinariedad parten, en principio, de la definición de lo que son las disciplinas (campos, materias, áreas temáticas o temas y frentes de investigación), para posteriormente ver sus relaciones, los solapamientos que se producen entre ellas, la mayor o menor cercanía que poseen entre sí y las transacciones o flujos entre unas disciplinas y otras, estudiando por ejemplo, las citas entre conjuntos de títulos de revistas científicas entre sí (Leydesdorff y Cozzens, 1993).

Otra aproximación del enfoque bibliométrico al fenómeno, consiste en el estudio de las relaciones estructurales entre las disciplinas a partir de la construcción de mapas de ciencia basados en citas o cocitas, redes de autores y de instituciones que trabajan en una disciplina o mapas construidos a partir de los diversos contenidos léxicos que integran los documentos de las bases de datos bibliográficas (McCain, 1986; Tijssen, 1992; Zitt, 2005).

Los mapas de cocitación pueden realizarse con diferentes niveles de agregación, dependiendo tanto de las unidades empleadas como del ámbito del estudio y modelan los distintos dominios del conocimiento, bien desde perspectivas macro, porque mapean la estructura global de la ciencia (Vargas Quesada, 2005), a nivel nacional, a partir de datos de producción de un país (Moya Anegón, et. al., 2006) o de nuevos frentes de investigación emergentes (McCain, 1998). De ese modo, se revelan las estructuras cognitivas e intelectuales de las disciplinas, aparecen las dimensiones, las posiciones y los patrones de comunicación científica formal y las redes intelectuales que integran las diferentes especialidades en las que trabajan los científicos (White, 1990). Con una aproximación de este tipo, se puede responder en cierta medida a la pregunta acerca del grado en que las disciplinas son multidisciplinarias o interdisciplinarias e identificar cuál es el grado de dependencia mutua entre ellas y la amplitud de contextos con los que se trabaja en un dominio temático.

Ahora bien, las actividades de comunicación científica que miden los estudios bibliométricos, son también procesos sociales, donde confluyen actores con diferentes perspectivas disciplinares, insertos en instituciones heterogéneas desde el punto de vista intelectual y organizativo, en el que conviven equipos de investigación donde se combinan habilidades y saberes diversos. El estudio de los patrones organizativos y de las interconexiones entre las organizaciones que se producen en un campo científico, mediante técnicas bibliométricas, ayuda a comprender e ilustrar la organización social que controla y dirige la investigación en un determinado dominio. Se puede apreciar si la organización basada en disciplinas universitarias sigue siendo el modo prevalente de unidad de producción del conocimiento. La hipótesis que aquí se mantiene, es que las disciplinas académicas entendidas como las unidades sociales de organización del conocimiento, son una variable histórica y, en consecuencia, no son una característica esencial de la ciencia contemporánea.

En esta comunicación se presentan los resultados de un estudio bibliométrico sobre las relaciones estructurales en la disciplina de la Biología Molecular, Celular y Genética a partir de la producción científica en este campo en la Comunidad de Madrid, en el período comprendido entre 1995 y 2003, mediante un análisis de cocitación de categorías ISI-JCR y de colaboración científica institucional, empleando técnicas de redes sociales. El objetivo del trabajo, ha sido examinar las relaciones disciplinares e institucionales que se producen en este campo. Para ello, se han construido mapas heliocéntricos de cocitación de categorías,

mediante técnicas que descubren la estructura de la red intelectual de la disciplina y mapas de colaboración institucional para explorar la red social de instituciones que colaboran en la producción científica en este dominio.

## 2 Material y Métodos

En un trabajo anterior (Olmeda Gómez, et. al., 2006) se ha detallado la procedencia y tratamiento de los datos, que ahora aquí se resumen. El 15 de enero de 2006 se descargaron los registros del *Web of Science (The Thomson Corporation)* correspondientes a los años 1995 a 2003 de todos los documentos publicados en España contenidos en las bases *Science Citation Index Expanded*, *Social Science Citation Index* y *Arts and Humanities Citation Index*, que contuvieran el término “Madrid”, en el campo *Address*. Los 68077 registros descargados de todo tipo de documentos (*articles, meeting abstracts, letters, reviews, book reviews notes, corrections, editorial materials* y *news items*), fueron exportados una vez normalizados, a una base de datos relacional que refleja la información estructurada de los documentos en los que queda asociada a cada uno de ellos, la categoría temática con la que *Thomson Scientific* asigna a cada una de las revistas en las que se han publicado.

Uno de los problemas que surgen cuando se quieren llevar a cabo análisis bibliométricos sobre las disciplinas científicas, es la forma en la que se asignan los documentos a las diferentes áreas científicas. La única información práctica que se tiene sobre la adscripción temática de los documentos a un campo científico determinado con los datos de procedencia, viene dada por las categorías temáticas (*Subject Category*) en las que el ISI divide el conocimiento científico en el *Journal Citation Report* y en la que están adscritas cada una de las revistas. El uso de los esquemas de clasificación de las revistas para categorizar los artículos en categorías temáticas es una técnica bibliométrica bien fundamentada y ha sido propuesta y experimentada como unidad de cocitación para la visualización de dominios científicos específicos (Moya Anegón, et. al., 2004; Vargas Quesada, 2005). Aunque no es un esquema de clasificación perfecto, tiene la ventaja de haber sido normalizado durante muchos años, es económico para el investigador desde el punto de vista de cálculo y se usa internacionalmente. Además, desde el momento en que las revistas se asignan a una o más categorías, se pueden elaborar, al menos, un mínimo de indicadores que permiten explorar las actividades en áreas multidisciplinares de Ciencia y Tecnología.

Mediante los procedimientos descritos, los documentos adscritos a la clase Biología Molecular Celular y Genética, proceden de los publicados en las revistas clasificadas en las siguientes categorías ISI:

CLASE	CATEGORÍA ISI
BIOLOGÍA MOLECULAR, CELULAR Y GENÉTICA (MOL)	Anatomía y Morfología
	Bioquímica y Biología Molecular
	Biología
	Biotecnología y Microbiología Aplicada
	Métodos de Investigación Bioquímica
	Biología Celular
	Virología
	Microbiología
	Biofísica
	Genética y Herencia
	Inmunología
	Toxicología
	Biología, Miscelánea
	Biología del Desarrollo
	Microscopía
Biología Evolutiva	

Tabla 1. Correspondencia de la clase Biología Molecular, Celular y Genética con las categorías ISI

El estudio tiene dos elementos principales: a) la construcción de mapas heliocéntricos mediante el análisis de cocitación de categorías y b) la construcción de mapas de colaboración institucional a nivel nacional en la misma clase.

El análisis de cocitación ha sido realizado con las técnicas descritas por Moya et. al. (2005, 2006). De forma muy resumida consiste en: el uso de valores no latentes de cocitación de categorías como medida de la similaridad, para cuantificar la relación existentes entre los 13179 documentos durante el período en Madrid, que han cocitado trabajos publicados en revistas pertenecientes a cualquiera de categorías del JCR adscritas a la clase de Biología Molecular, Celular y Genética; utilización de valores puros de cocitación sin normalizar; uso de las categorías ISI como elementos de representación y aplicación del algoritmo Kamada and Kawai (1989) para generar los grafos, fijando el nodo más productivo de la clase en la posición central del mismo. Los grafos resultantes se exportan al formato vectorial escalar SVG.

Los mapas de colaboración institucional, se han construido a partir de los datos brutos de coautoría interinstitucional, mediante recuentos completos. Un documento se considera coautorado sólo si sus autores tienen diferentes afiliaciones institucionales nacionales y uno de ellos, al menos, es de Madrid. Se han contabilizado 13399 colaboraciones correspondientes a la producción madrileña en la clase temática estudiada. Al igual que en los mapas heliocéntricos, se ha usado el algoritmo Kamada Kawai para generar los grafos, en este caso, sin fijar ningún nodo en el centro de la representación, pero sí adscribiendo con colores la pertenencia de cada institución a su respectiva Comunidad Autónoma.

La existencia de una colaboración entre dos instituciones implica que existe reciprocidad: si la institución X colabora con la institución Y, la 2ª colabora con la primera, pero no indica el grado de dependencia de una sobre otra.

Este grado de dependencia puede variar entre las distintas instituciones, ya que los grados de colaboración pueden que no sean simétricos. Para representar esta desemejanza en el grado de colaboración entre instituciones, se ha calculado la tasa de colaboración asimétrica (TCA), entre las instituciones que integran la red.

La idea de este indicador procede del Índice de Afinidad empleado para los cálculos de relaciones asimétricas entre dos países (Zitt, 2000), pero adaptándolo para el cálculo entre la colaboración asimétrica entre instituciones. Se ha calculado del modo siguiente, con dos fórmulas que miden el sentido de la colaboración entre dos nodos cualesquiera:

$$TCA(Insti_1 \rightarrow Insti_2) = \frac{COL(Insti_1 \leftrightarrow Insti_2)}{COL(Insti_1 \leftrightarrow total_{clase})} \times 100$$

$$TCA(Insti_2 \rightarrow Insti_1) = \frac{COL(Insti_2 \leftrightarrow Insti_1)}{COL(Insti_2 \leftrightarrow total_{clase})} \times 100$$

Fig. 1 Fórmulas de las tasas de colaboración asimétrica

En el mapa construido al efecto, el sentido de la flecha indica la dependencia de la institución en cuestión respecto con quien se vincula.

### 3 Resultados y Discusión

La figura 2, representa el mapa heliocéntrico de cocitación de categorías ISI de la producción madrileña en la clase de Biología Molecular, Celular y Genética, representándose sólo las 16 categorías ISI en las que se descompone esta clase temática, según se ha señalado en la tabla 1. Cada nodo tiene un tamaño variable dependiendo del volumen de documentos cocitados perteneciente a la categoría con la que está etiquetado y el título se corresponde con la denominación que el JCR adjudica a la revista en la que se ha publicado el trabajo.

Se ha elegido como categoría central, aquella que cuenta con mayor número de trabajos cocitados en la clase, que es la Bioquímica y Biología Molecular. Las líneas que conectan los nodos representan los valores de cocitación entre las categorías del dominio y su longitud variable viene expresada por los valores de cocitación. Los valores más altos son transformados en vínculos de menor longitud desde el centro, es decir con mayor similitud y los valores más bajos de cocitación, que expresan mayor disimilitud, con longitudes más largas. En las órbitas más cercanas orbitan, en consecuencia, aquellas categorías dentro de su clase con las que la central tiene mayor similitud, a partir de los vínculos intelectuales que emergen de los valores de cocitación que han construido los científicos que trabajan en Madrid.

En el mapa aparecen dos zonas diferenciadas. La primera se corresponde con las categorías más cercanas, las más similares, cuyo orden atendiendo al grado de similitud es el siguiente: Biología Celular, Biofísica y Métodos de Investigación Bioquímica, Genética y Herencia, Inmunología y Biología. Las trayectorias más alejadas son ocupadas por Microbiología, Biología del Desarrollo, Virología, Toxicología, Biología-Miscelánea, Biología Evolutiva y Microscopía. Ahora bien, los mapas de cocitación son similares a las fotos instantáneas o cortes transversales de una red en continua transformación (Small, 2000) y su potencialidad para visualizar el carácter multidisciplinar de la estructura de un dominio científico reside en esa particularidad.

Es muy probable que como consecuencia del grado multidisciplinar de la clase, existan más “planetas”, que los representados en el mapa de la Figura 2 y que pueden ser importantes al analizar el dominio temático. De hecho, cada una de los 16 nodos de la Fig. 2, es susceptible

de convertirse en el nodo central de la red heliocéntrica y dar origen a un nuevo mapa que incluya el conjunto de categorías con las que es más cocitada. Ante la imposibilidad de representar ese volumen de información (16 mapas de cocitación, uno por categoría) se ha optado como ejemplo, representar el mapa de la categoría correspondiente a la Bioquímica y Biología Molecular (Fig.3).

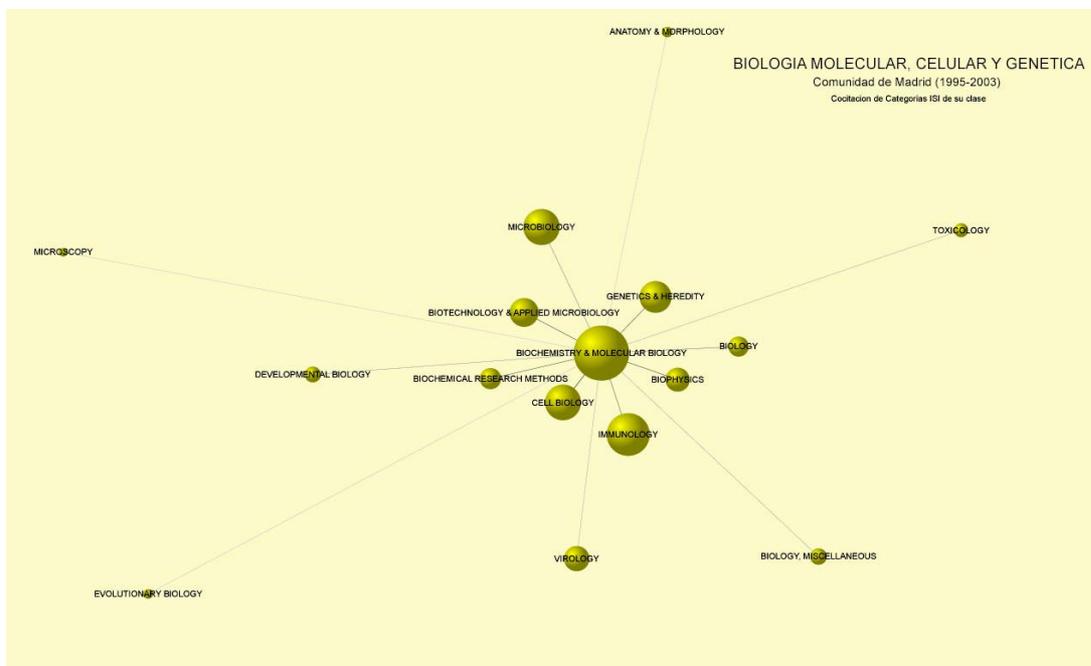


Fig 2. Mapa heliocéntrico de la Biología Molecular, Celular y Genética de la Comunidad de Madrid. 1995-2003. Cocitación de categorías sólo en su clase.

Este mapa heliocéntrico representa el mapa de cocitación de la misma clase, la Biología Molecular, Celular y Genética con las categorías con las que es más cocitada, independientemente de la clase a la que pertenezca.

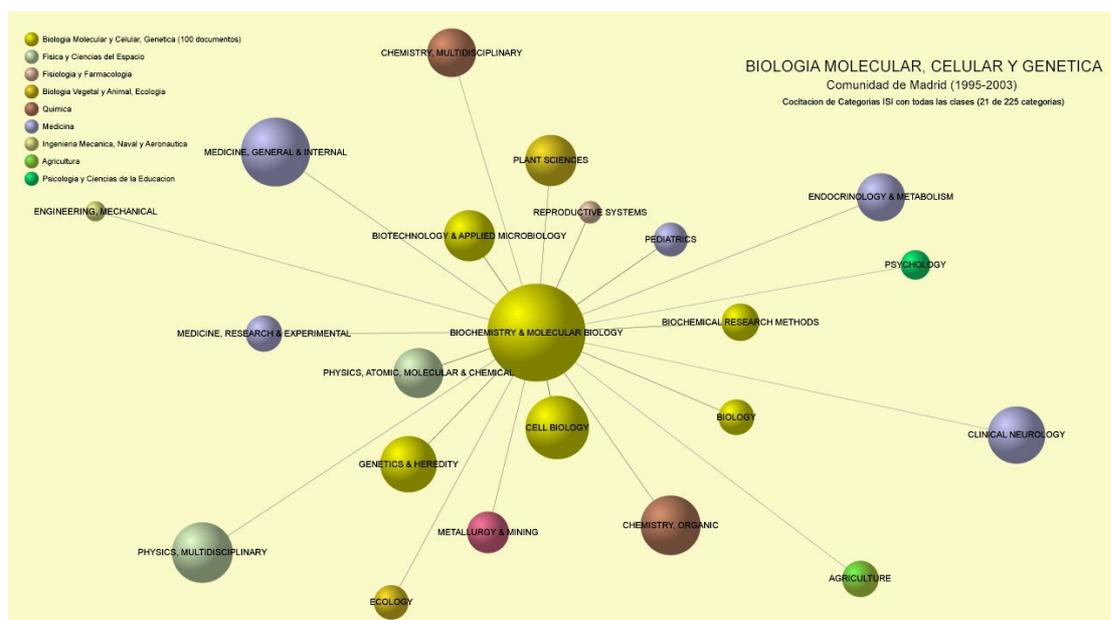


Fig. 3. Mapa heliocéntrico de la Biología Molecular, Celular y Genética de la Comunidad de Madrid. 1995-2003. Cocitación de categorías ISI con todas las clases. 21 de 225 categorías.

En este mapa y, a diferencia del anterior, aparecen 16 categorías nuevas, que a su vez pertenecen a 8 clases diferentes, cuyos colores aparecen en la leyenda del mapa. Lógicamente, los nodos que pertenecen a estas clases orbitan a mayor distancia que las que pertenecen a su clase, y por ello se sitúan en órbitas exteriores. Al ser sólo las 21 primeras de 225, se visualiza el grado de interconexión y las categorías con las que más se interrelaciona. Además se aprecia que el centro de esta categoría está ocupado por nodos en su mayoría, de su misma clase, con la excepción de la Física Atómica, Molecular y Química, lo que puede indicar que el dominio cuenta con un centro estable de subespecialidades.

Por otra parte, si se entiende que el grado de “heliocentricidad” de un dominio es directamente proporcional a su nivel multidisciplinar, se puede apreciar el elevado grado de esta cualidad en la categoría elegida en la representación, ya que cocita en total con 225 categorías. Este conjunto de categorías ajenas a su clase que orbitan a mayor distancia, vienen a ilustrar el grado y la extensión de la segmentación de campos, de la diversidad de fines y de enfoques conceptuales en el dominio.

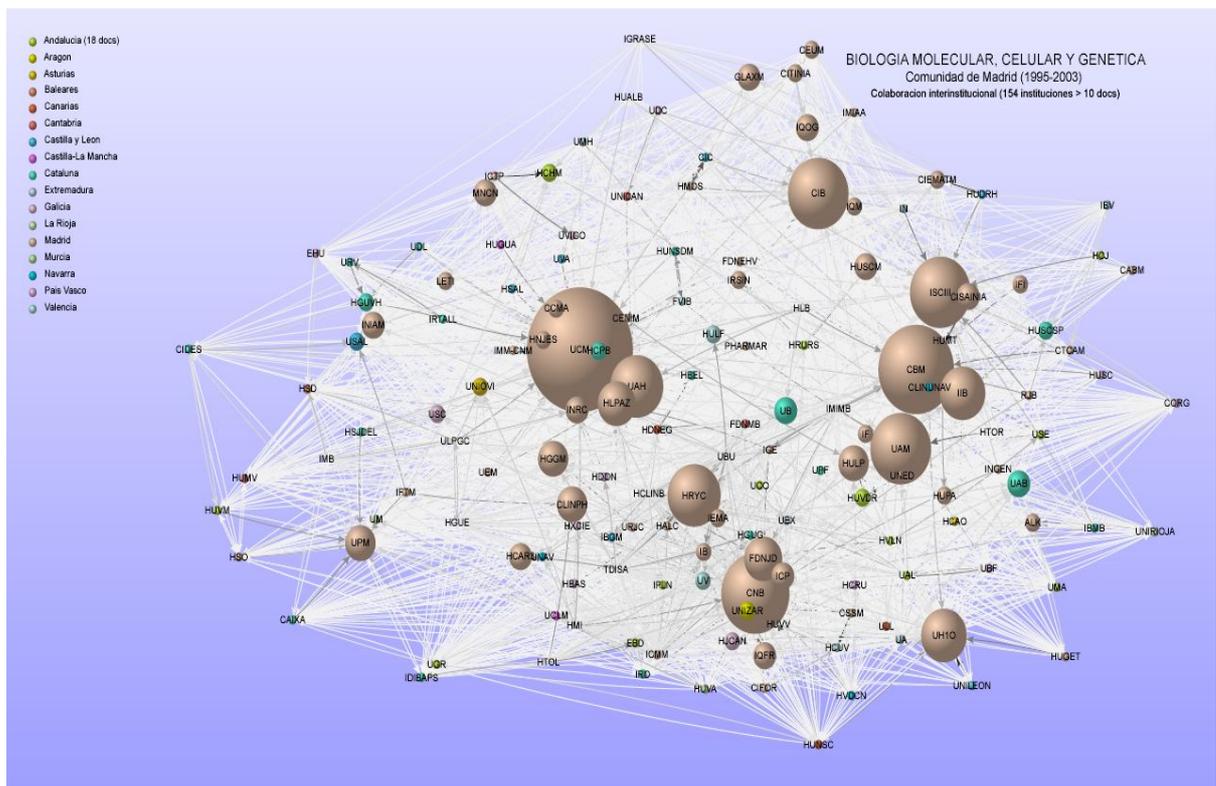


Fig. 4. Mapa de colaboración asimétrica. Biología Molecular, Celular y Genética. CAM 1995-2003. Colaboración interinstitucional (154 instituciones > 10 documentos)

La colaboración interinstitucional ha sido representada tal y como aparece en la Figura 4. Los nodos cuentan con una producción en colaboración mínima superior a 10 documentos y con estos valores aparecen 154 instituciones en el mapa, ya que al disminuir o aumentar ese valor bruto de documentos en colaboración, la red se hace más o menos densa. Aparecen los grandes nodos productores que sirven de aglutinadores de 4 grandes clusters. Por tamaños de producción en colaboración, el primer cluster corresponde al formado por la Universidad Complutense, la Universidad de Alcalá de Henares, el Hospital La Paz, el Instituto de Neurobiología Ramón y Cajal y el Hospital General Gregorio Marañón. Siguiendo el

movimiento de las agujas del reloj y en la parte superior derecha del grafo aparece el cluster del Centro Nacional de Biología, con vínculos con instituciones tales como el Instituto Químico Médico, El Instituto de Química Orgánica General, la Universidad San Pablo CEU, la empresa Glaxo. A continuación, aparece un macro cluster que es el resultado de la colaboración intensa entre 4 grandes centros de producción, como son el Centro de Biología Molecular Severo Ochoa, la Universidad Autónoma, el Instituto de Salud Carlos III, el Instituto de Biotecnología, se sitúan la Fundación Jiménez Díaz, el Instituto de Catálisis y Petroquímica y el Instituto Químico-Físico Rocasolano.

## 4 Conclusiones

El análisis conduce a las siguientes conclusiones: las técnicas de cocitación de categorías ISI-JCR permiten una representación válida para el análisis de dominios y subdominios regionales. Es una herramienta útil para identificar y analizar en dominios multidisciplinares regionales como la Biología Molecular, Celular y Genética, el grado de interconexión con otras especialidades. Permite distinguir las especialidades centrales o subespecialidades, de otras con las que mantiene contactos. La aplicación de técnicas similares para la construcción de redes de coautoría, posibilita identificar los actores e instituciones que producen en dominios específicos. La utilización de diferentes valores de corte que varían los niveles de agregación posibilita su uso para el estudio de frentes de investigación y las técnicas, en conjunto, pueden ser incorporadas para enriquecer las disponibles en el campo de la “espectrometría bibliométrica”.

\*

Este trabajo ha sido posible gracias a la financiación concedida por la Comunidad de Madrid en el proyecto 06/HSE/0166/2004

## Bibliografía citada

- KAMADA, T.; KAWAI, S. An algorithm for drawing general undirected graphs. *Information Processing Letters*, 1989, vol. 31, n. 1, p. 7-15.
- LEYDESDORFF, L.; COZZENS, S. E. The delineation of specialties in terms of journals using the dynamic journal set of the Science Citation Index. *Scientometrics*, 1993, vol. 26, n. 1, p.133-154.
- McCAIN, K. W. Cocited author mapping as a valid representation of intellectual structure. *Journal of the American Society for Information Science*, 1986, vol. 37, n. 3, p. 111-122.
- McCAIN, K. W. Neural networks research in context: a longitudinal journal cocitation analysis of an emerging interdisciplinary field. *Scientometrics*, 1998, vol. 41, n. 3, p.389-410.
- MOYA ANEGÓN, F. [et al.]. A new technique for building maps of large scientific domains on the cocitation of classes and categories. *Scientometrics*, 2004, vol. 61, n.1, p. 129-145.
- MOYA ANEGÓN, F. [et al.]. Domain analysis and information retrieval through the construction of heliocentric maps based on ISI-JCR category cocitation. *Information Processing and Management*, 2005, vol. 41, p. 1520-1533.

- MOYA ANEGÓN, F. [et al.]. Visualización de la estructura científica española: 1990-2005. *El Profesional de la Información*, 2006, vol. 15, n. 4, p. 258-269.
- OLMEDA GÓMEZ, C. [et al.]. *Indicadores científicos de Madrid (ISI, Web of Science, 1990-2003)*. Madrid: Dirección General de Universidades e Investigación, 2006.
- SMALL, H. Charting pathways through science: exploring Garfield's vision of a unified index to science. En: CRONIN, B.; ATKINS, H. B. (eds.). *The web of Knowledge: a Festschrift in honor of Eugene Garfield*. Medford: Information Today, 2000. p. 449-473.
- TIJSEN, R. J. W. A quantitative assessment of interdisciplinary structures in science and technology: co-classification analysis of energy research. *Research Policy*, 1992, vol. 21, n. 1, p. 27-45.
- VARGAS-QUESADA, B. *Visualización y análisis de grandes dominios científicos mediante redes pathfinder*. Tesis doctoral. Granada: Universidad, 2005.
- WHITE, H. D. Author cocitation analysis: overview and defense. En: BORGMAN, C. (ed.). *Bibliometric and scholarly communication*. Newbury Park, CA: Sage, 1990. p. 84-106.
- ZITT, M. Facing diversity of science: a challenge for bibliometric indicators. *Measurement*, 2005, vol.3, n.1, p. 38-49.
- ZITT, M.; BASSECOULARD, E.; OKUBO, Y. Shadows of the past in international cooperation. Collaboration profiles of the top five producers of science. *Scientometrics*, 2000, vol. 47, n. 3, p. 627-657.