

# Visualización y análisis de la estructura científica española: *ISI Web of science* 1990–2005

Por Félix de Moya–Anegón, Benjamín Vargas–Quesada, Zaida Chinchilla–Rodríguez, Elena Corera–Álvarez, Antonio González–Molina, Francisco José Muñoz–Fernández y Víctor Herrero–Solana

**Resumen:** En este trabajo presentamos la estructura de la investigación científica española durante el período comprendido entre 1990 y 2005. Basándonos en la cocitación de categorías ISI, que comprenden las ciencias naturales, las ciencias sociales y el arte y las humanidades, y combinando el análisis de redes sociales (ARS) con algoritmos de poda como *pathfindernetorks* (Pfnets) y otros algoritmos de representación de la información como el de Kamada–Kawai, mostramos la estructura básica de la ciencia española a través de sus relaciones esenciales. El resultado son unos mapas bibliométricos con forma de neurona y con un gran axón central (cienciogramas) que ponen de manifiesto la macro–estructura, micro–estructura y columna vertebral de la investigación española, que permiten su análisis e interpretación.

**Palabras clave:** Visualización de información, Análisis de dominios, Producción científica, Cocitación, España.



Félix de Moya–Anegón, catedrático del área de biblioteconomía y documentación de la Universidad de Granada y responsable del Grupo de Investigación SCImago dedicado al análisis, visualización y recuperación de la información científica y que desarrolla el proyecto del Plan Nacional de I+D+I Atlas de la Ciencia.

Benjamín Vargas–Quesada, doctor en documentación e información científica. Profesor de la Facultad de Documentación de la Universidad de Alcalá e investigador del Grupo SCImago.

Zaida Chinchilla–Rodríguez, doctora en documentación e información científica. Investigadora del Grupo SCImago.

Elena Corera–Álvarez, licenciada en documentación. Becaria predoctoral de la Universidad de Granada e investigadora del Grupo SCImago.

Antonio González–Molina, licenciado en documentación. Investigador del Grupo SCImago.

Francisco José Muñoz–Fernández, licenciado en documentación. Becario de investigación del programa de Formación de Personal Docente de Investigador de la Junta de Andalucía, e investigador del Grupo SCImago.

Víctor Herrero–Solana, doctor en documentación. Profesor de la Facultad de Documentación de la Universidad de Granada e investigador del Grupo SCImago.

**Title:** Visualizing and analyzing the Spanish science structure: *ISI Web of science* 1990–2005

**Abstract:** This article presents the structure of Spanish science from 1990 to 2005. Based on ISI subject–category co–citation, including natural sciences, social sciences and arts & humanities, and combining both social network analysis (SNA) and pruning algorithms such as *pathfinder–networks* (Pfnets) and other information layout algorithms like Kamada–Kawai's, the basic structure of Spanish science is shown through its essential relations. The results are bibliometric–maps (scientograms) revealing the macrostructure, microstructure and the backbone of Spanish science and making possible its analysis and explanation.

**Keywords:** Information visualization, Domain analysis, Scientific output, Cocitation, Spain.

Moya–Anegón, Félix de; Vargas–Quesada, Benjamín; Chinchilla–Rodríguez, Zaida; Corera–Álvarez, Elena; González–Molina, Antonio; Muñoz–Fernández, Francisco José; Herrero–Solana, Víctor. «Visualización y análisis de la estructura científica española: *ISI Web of science* 1990–2005». En: *El profesional de la información*, 2006, julio–agosto, v. 15, n. 4, pp. 258–269.

## 1. Introducción

La representación gráfica de la información para su posterior visualización es una actividad común en la mayoría de las disciplinas científicas en los últimos tiempos (Klovdhal, 1981), (Crosby, 1997). Pero el uso de las representaciones gráficas, en combinación con la tecnología informática para conseguir una adecuada visualización de la información, es una tarea relativamente nueva que se ha convertido en uno de los principales objetos de estudio.

La visualización de la información no es el resultado implícito del acto de ver, no es un producto espontáneo del individuo que recibe la información ya visualizada. Es una tarea del comunicador, que transforma datos abstractos y fenómenos complejos de la realidad en mensajes visibles, haciendo posible que los individuos vean con sus propios ojos datos y fenómenos que son directamente inaprensibles y que, por tanto, comprendan la información que yace oculta (Costa, 1998). El dominio de los lenguajes gráficos para visualizar estos “efectos invisibles” configura una nueva ciencia de la comunicación, la esquemática, a la que Costa ha definido como el tercer lenguaje, después de la imagen y el signo.

---

**«La visualización de información es el proceso de comunicación a partir de la cual es posible percibir hechos y fenómenos que de otra forma pasarían inadvertidos»**

---

Por otro lado, la visualización de la información consiste en un acto de transferencia de conocimiento. Se justifica en el hecho de que el mundo es multifacético, multidimensional, multifenomenico y se presenta como un “continuum”. Por tanto, hay que conseguir la integración de la imagen, la palabra, el número y el arte. Los instrumentos necesarios para conseguir este

objetivo son aquellos relacionados con la escritura y la tipografía, la gestión de grandes cantidades de datos y el análisis estadístico, los gráficos, su distribución y el color. Y los estándares de calidad imprescindibles se derivan de los principios visuales que nos dicen qué poner en el sitio adecuado (Tufte, 1994). La dificultad surge en el momento de transferir la complejidad del mundo, su dinamismo y multidimensionalidad, a un formato comprimido, reducido, estático y plano o sin relieve, como es el papel o la pantalla de un ordenador. Para contrarrestar estas limitaciones, Tufte propone seis ideas:

—Evitar en todo momento las representaciones planas y recurrir a los gráficos tridimensionales o al menos bidimensionales con perspectiva (2,5D).

—Realizar micro y macro representaciones de una misma realidad.

—Evitar los agrupamientos desordenados de elementos, descubriendo los detalles y elementos complejos por medio de la distribución adecuada de la información y su visualización estratificada.

—Conseguir distintas imágenes de una misma realidad desde diferentes perspectivas.

—Asociar colores a la información.

—Conseguir gráficos en cuatro dimensiones: combinación de imágenes tridimensionales con información temporal relativa a las mismas.

Fusionando las ideas de Costa y Tufte podemos decir que la visualización es el proceso de comunicación que se produce entre una representación reducida de la realidad y quien la observa, a partir de la cual es posible percibir por medio de la vista, hechos y fenómenos de la realidad multidimensional y cambiante del mundo que de otra forma pasarían totalmente inadvertidos. Siguiendo con la fusión de ideas, diremos que la visualización de la información consiste en la intersección de la imagen, la palabra, el número y el arte, a través de la escritura y la tipografía, la gestión

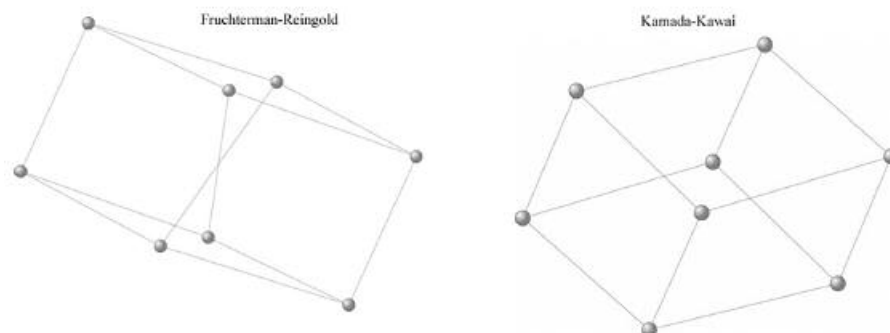


Figura 1

de grandes cantidades de datos y el análisis estadístico, los gráficos, su distribución y el color.

No es una práctica nueva en el campo de la documentación. Incluso se podría considerar a la visualización de la información científica como a una de las pioneras en estas lides. Sugerida hace más sesenta años por **Bush** en 1945 y puesta en práctica hace ahora poco más de cuarenta años por **Garfield, Sher** y **Torpie** en 1964, ha sido utilizada para poner al descubierto y divulgar la esencia y estructura de la ciencia. Pese a sus años de existencia, aún hoy día se encuentra en fase de evolución y desarrollo. Prueba de ello es su aplicación en nuevos campos o actividades, como por ejemplo la visualización y análisis de dominios científicos.

Los mapas bibliográficos reflejan las relaciones entre disciplinas, y sus etiquetas revelan las conexiones semánticas al tiempo que sirven de indicio para conocer por qué unas están conectadas con otras. Es más, ponen de manifiesto qué dominios de la ciencia o especialidades son objeto de investigación hoy día, y qué individuos, publicaciones, instituciones, regiones o países son los más prominentes (**Garfield, 1994**). A la construcción de mapas a partir de información bibliométrica también se le conoce como *cienciografía* y,

aunque es un término poco utilizado, posiblemente debido a la proliferación de otros como visualización de dominios o dominios de conocimiento que están relacionados con la misma idea, en nuestra opinión se trata de una locución muy adecuada para referirse al hecho de cartografiar la ciencia mediante grafos.

«Las mejores visualizaciones son las que se obtienen utilizando cocitación en su estado puro, es decir, sin normalizar»

Éste es el primero de una serie de trabajos destinados a la visualización, análisis y estudio de la evolución del dominio científico español mediante *cienciogramas*, que son el resultado de la *cienciografía*.

A continuación realizamos una breve revisión de los trabajos publicados en este campo. Seguimos con el desarrollo metodológico y su validación. Después mostramos los resultados, procediendo a su análisis en tres niveles: macro-estructura, micro-estructura y columna vertebral. Continuamos con la discusión para finalizar con unas breves conclusiones.

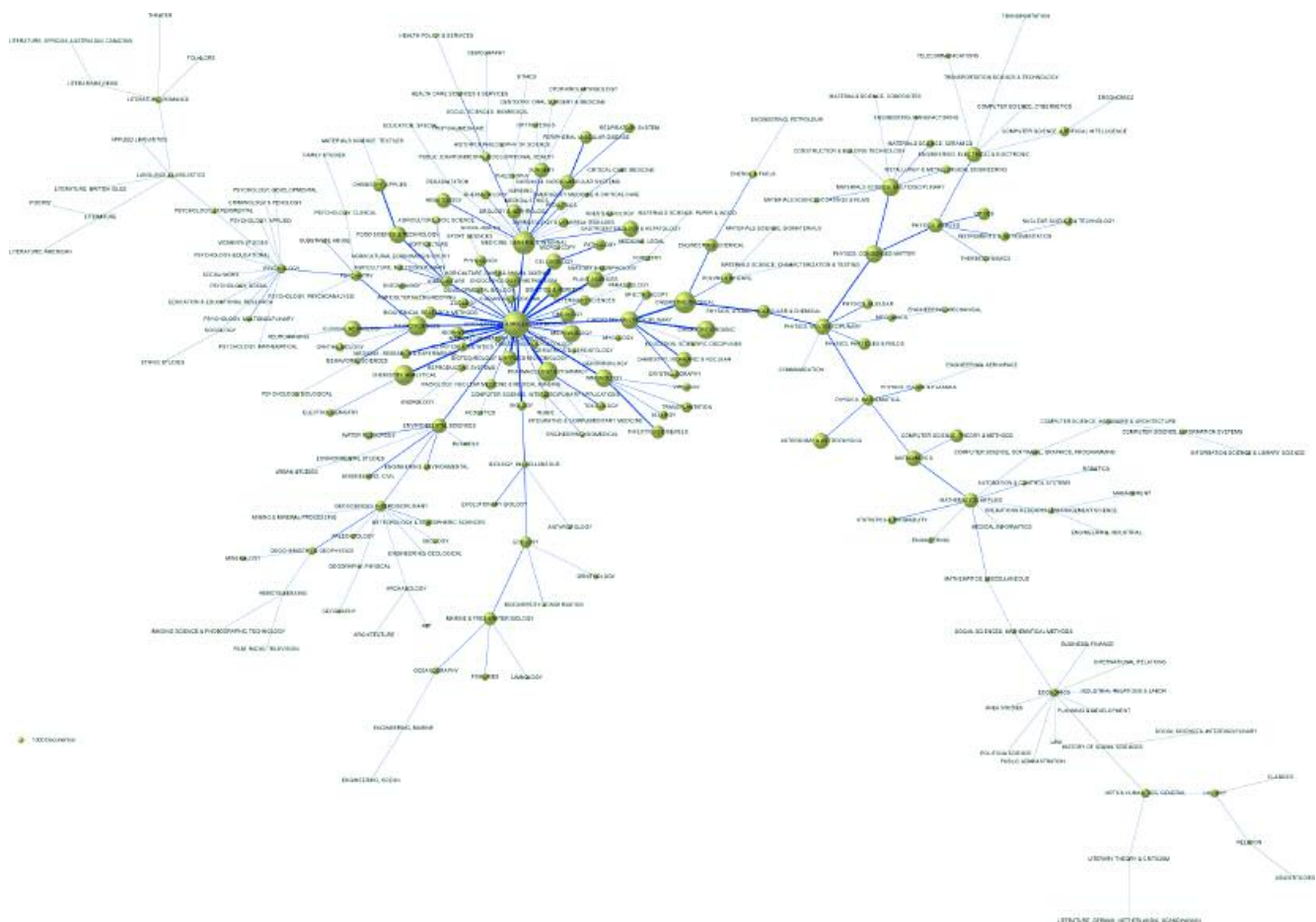


Figura 2. Cienciograma de la producción científica española, 1990–2005

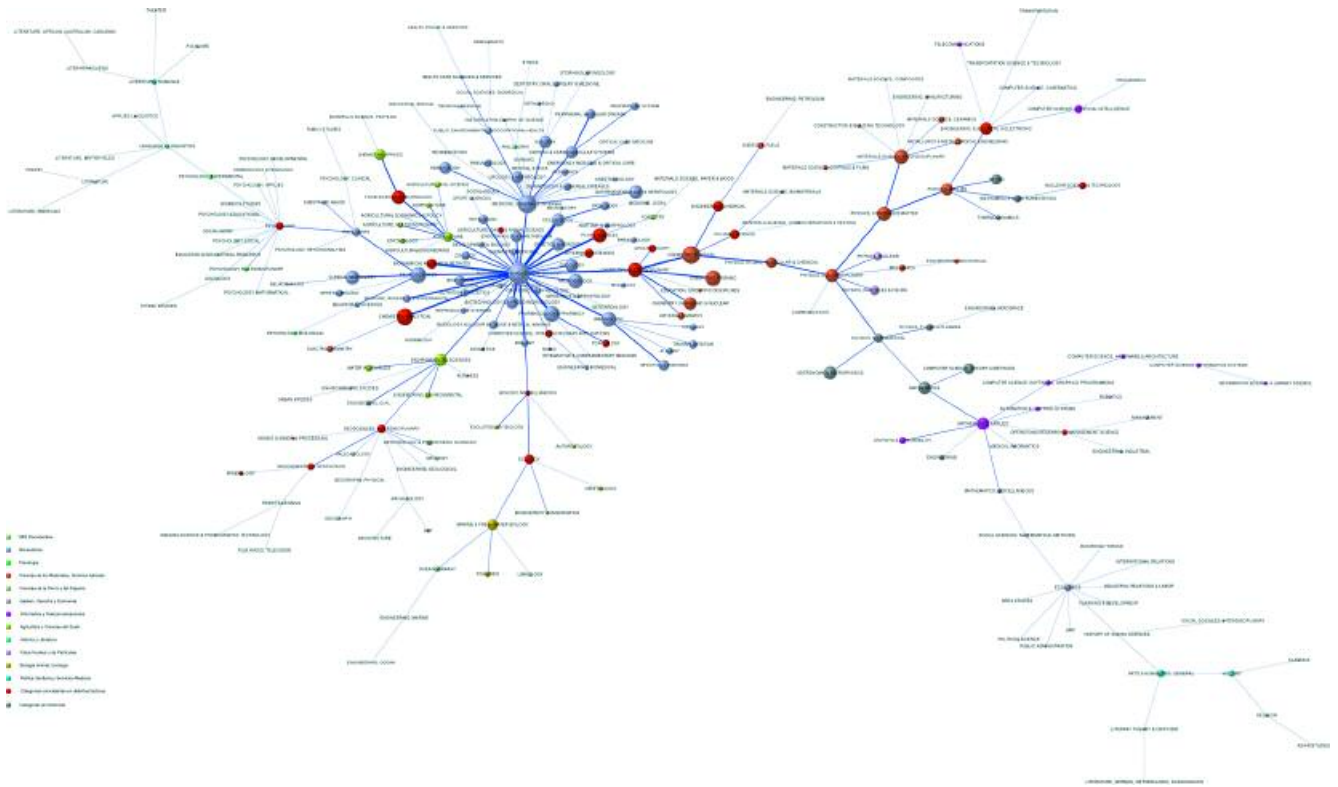


Figura 3. Cienciograma factorial de España 1990–2005

## 2. Antecedentes

En Moya–Aneón et al., (2004), revisamos histórica y profundamente la evolución de los mapas de la ciencia desde su origen hasta nuestros días, propusimos la cocitación de categorías *ISI–Journal Citation Report (JCR) (The Thomson Corporation, 2005)* para la representación de grandes dominios científicos, y demostramos su utilidad mediante la visualización del dominio geográfico español en el año 2000. Desde entonces hasta ahora han aparecido algunos trabajos relacionados con la visualización de grandes dominios científicos, pero salvo uno basado en la cocitación de categorías, Moya–Aneón et al.(2005), el resto utilizan revistas como unidad de análisis. Leydesdorff (2004a,b) clasifica la ciencia mundial mediante el algoritmo gráfico analítico de componentes biconectados, en combinación con el *JCR 2001*. Boyack, Klavans y Börner (2005) aplican ocho medidas diferentes de similitud a una base de 7.121 revistas procedentes del *Science Citation Index (SCI)* y del *Social*

*SCI*, que aglutinan alrededor de un millón de documentos, mostrando lo que ellos han dado en llamar el primer mapa global de la ciencia. Samoylenko, et al., (en prensa) proponen una aproximación a la visualización de la ciencia mundial y su evolución mediante *Minimum spanning trees (MST)* de las revistas del *SCI* durante el período 1994–2001.

Nosotros, en este trabajo generaremos cienciogramas basados en la cocitación de categorías *ISI–JCR* podadas mediante *Pathfindernetworks (Pfnets)* (Schvaneveldt, 1990), representadas mediante el algoritmo de Kamada–Kawai, 1989, y aplicadas sobre la producción científica mundial de las bases de datos *SCI–Expanded*, *SSCI* y *Arts & Humanities Citation Index (A&HCI)*.

## 3. Metodología

Para poner de manifiesto la estructura científica del dominio español, aplicamos un método diseñado para la visualización de grandes dominios científicos, explicado y desarrollado en Vargas–Quesada (2005) y en Moya–Aneón et al. (en prensa) que sigue el proceso propuesto por Börner, Chen y Boyack en 2003.

### 3.1. Recopilación de información

Con fines estrictamente investigadores, el 15 de enero de 2006, terminamos de descargar del *Web of science (The Thomson Corporation, 2006a)*, en concreto del *SCI–Expanded*, *SSCI* y *A&HCI*, todos los registros de la producción científica española publicados

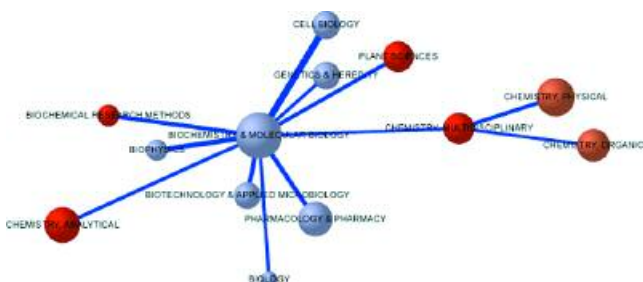


Figura 4. Columna vertebral de la investigación española 1990–2005

entre los años 1990 y 2005. El *ISI* (actualmente registrado como *Thomson Scientific*) asigna a cada una de las revistas una o varias categorías temáticas. Con el fin de poder adjudicar una materia (categoría *ISI*) a cada documento, disponemos del *JCR* tanto en su edición *Science* como *Social Science*, para los años que se pretenden cubrir. Los registros descargados fueron exportados a una base de datos relacional que refleja la información estructurada de los documentos. Este nuevo repositorio contiene 372.002 documentos fuente (*articles, biographical items, book reviews, corrections, editorial materials, letters, meeting abstracts, news items y reviews*) que se han publicado en 7.585 revistas *ISI* clasificadas en 241 categorías, que citan a 9.550.959 documentos *ISI* anteriormente publicados.

### 3.2. Unidades de análisis

Las más comúnmente utilizadas para la representación de dominios científicos son: revistas, documentos, autores, términos y palabras. Recientemente a este grupo se les han añadido también otras más amplias como son: países, ámbitos temáticos de diferentes niveles, instituciones y categorías *ISI*. A pesar de que hay autores que no comparten estas ideas (**Leydesdorff**, en prensa), consideramos que las mencionadas categorías *ISI* son unidades informativas lo suficientemente explícitas como para poder ser utilizadas en la representación de las distintas disciplinas que componen la ciencia en general. En combinación con las adecuadas técnicas de reducción del espacio y de representación de la información, permiten la construcción de cienciogramas de la ciencia o de grandes dominios científicos, mucho más informativos y amigables, que los que se obtienen mediante la cocitación de unidades de análisis menores. Por todo ello, utilizaremos como unidades de medida las categorías productivas del dominio científico español, que recoge el *JCR* durante el período 1990–2005, con la excepción de *multidisciplinary sciences*. Una de las características del *JCR* es que adjudica esta categoría a un determinado grupo de revistas de naturaleza multidisciplinar como *Science, Nature, Endeavour* o *Interciencia*, entre otras. Esto, que en un primer momento parece lo más lógico y acertado, puesto que dan cabida a documentos de todo tipo de disciplinas, no lo es tanto si analizamos de forma detallada las consecuencias que esa decisión puede tener. Trabajos que tratan de una determinada materia o categoría, como podría ser por ejemplo la genética, aparecen desgajados del resto de sus similares al haber sido publicados en una revista multidisciplinar y ser etiquetados por tanto como tales. Este problema no es fácil de resolver. La *Comisión Europea*, en el *Third European report on science & technology indicators 2003: towards a knowledge-based economy* la ha eliminado directamente y, por tanto, los documentos al-

bergados bajo ella. En este trabajo, para evitar pérdidas de información de este tipo, sustituimos la categoría de los documentos que la tienen asignada por aquella más citada por las referencias de cada uno de ellos. Sólo en el caso de que *multidisciplinary sciences* sea la única citada por los documentos referenciados, hemos recuperado los textos fuente de esas referencias y, a partir de su título y resumen, le hemos asignado de forma manual una categoría *JCR*. El número máximo de unidades con el que trabajaremos en España será de 240.

### 3.3. Unidades de medida

De acuerdo con nuestra experiencia (**Moya-Aneón et al.**, 2005), utilizaremos la cocitación, como medida de similitud para cuantificar la relación existente entre cada una de las categorías del *JCR*. Tras muchos ensayos, hemos llegado a la conclusión de que las mejores visualizaciones son aquellas que se obtienen utilizando cocitación en su estado puro, es decir, sin normalizar. No obstante, y puesto que es necesario equiparar el grado de citación de las distintas bases de datos a los valores puros de cocitación obtenidos, les sumamos su correspondiente valor ya normalizado. De esta forma, conseguimos trabajar con cocitación pura, al mismo tiempo que normalizamos el grado de las distintas categorías. Esto lo logramos con una simple modificación de la ecuación para la normalización del grado de citación propuesta por **Salton y Bergmark** en 1979, donde:

$$MCN(ij) = Cc(ij) + \frac{Cc(ij)}{\sqrt{c(i) \cdot c(j)}}$$

Medida de Cocitación de **Salton y Bergmark** modificada

—*MCN*: medida de cocitación normalizada.

—*Cc*: Cocitación.

—*c*: citación.

—*i, j*: categorías

El resultado que se obtiene como consecuencia de aplicar esta medida es una matriz de cocitación simétrica de 240 por 240 categorías, la cual es la base y el origen de los cienciogramas que ponen de manifiesto la estructura del dominio que se representa, así como las relaciones o intercambio de información (conocimiento) que se produce entre las distintas categorías o disciplinas del dominio científico español.

### 3.4. Reducción y distribución de la información

#### 3.4.1. Reducción de la dimensión

A lo largo de la historia de la visualización de la información científica se han utilizado técnicas muy



distintas y variadas para reducir el espacio  $n$ -dimensional. Solas o en combinación con otras, las más comunes han sido: *Multidimensional scaling (MDS)*, *Clustering*, *Factor analysis*, *Self-organizing maps* (Kohonen, 1985; Kohonen, 1997) y *Pfnet*.

La adopción de las categorías *ISI* como unidades de medida implica la aceptación de dichos elementos como unidades de representación, lo cual inevitablemente conlleva a que los cienciogramas resultantes contengan normalmente un número superior a las doscientas categorías. Mostrar la imagen del dominio a partir de un número elevado de unidades fácilmente identificables por sus etiquetas, que muestren sus interacciones por medio de enlaces o relaciones, y hacerlo de una forma inteligible y estéticamente agradable para el ojo humano es una tarea difícil. Por ello, y teniendo en cuenta la advertencia de Hjörland y Albrechtsen (1995): «(...) si se trata de un sistema con demasiadas posibilidades, en el que no se le hayan dado prioridad a las conexiones esenciales, sólo se obtendrá un exceso de información por parte del usuario, y una falta de efectividad en lo que se refiere al sistema»; seguimos el consejo de Small (2000): (...) «por motivos de claridad en la visualización es mejor eliminar algunas conexiones (...) la pérdida de información de la estructura implica la ganancia en simplicidad, por lo que en algunos casos el sacrificio está justificado; y coincidimos plenamente con White (2003): «entre las distintas técnicas disponibles, las representaciones *Pfnet* basadas en cocitación pura y visualizadas mediante algoritmos del tipo *spring embedders*, parecen tener considerables ventajas».

---

**«Que information science & library science esté conectada con computer science information systems y no con otra, implica que los investigadores consideran que es la conexión más representativa del dominio español en el período de tiempo estudiado»**

---

Este proceso de simplificación de la ciencia no es nuevo. Se trata de un principio conocido como *Navaja de Occam*, que en un lenguaje cotidiano vendría a significar algo así como que hay que favorecer la hipótesis más sencilla que explique las observaciones. Esta máxima se ha llevado mucho más lejos, sobre todo desde que se empezó a admitir que las leyes físicas se escribían en el lenguaje de las matemáticas: la hipótesis más sencilla es, a priori, la que tiene una for-

mulación matemática más fácil. En el caso de la visualización del dominio científico español aplicamos este mismo principio minimalista: ¿por qué visualizar y analizar una estructura densa y compleja si podemos obtener y estudiar otra más sencilla y con las relaciones más significativas o esenciales? Consideramos que *Pfnet* con valores de poda  $r=\infty$ , y  $q=n-1$  (Chen, 1998), es la mejor técnica para eliminar las relaciones redundantes, preservando y poniendo de manifiesto las más significativas o esenciales de un dominio, al tiempo que captura de una forma económica su estructura intelectual.

### 3.4.2. Distribución espacial de la información

Existen muchos métodos para la generación automática de grafos. De todos ellos, los insertadores de muelles (*spring embedders*) son los más utilizados en el área de la documentación. Comienzan asignando coordenadas a los nodos de modo que el grafo final sea estéticamente agradable al ojo humano. A este proceso se le conoce como *embedding* (Eades, 1984). Las dos principales extensiones al algoritmo propuesto por Eades han sido desarrolladas por Kamada y Kawai (1989) y Fruchterman y Reingold (1991). Los criterios para evaluar este tipo de algoritmos son principalmente estéticos y, aunque estudios como el de Brandenburg, Himsolt y Rohrer en 1995 no detectan la preponderancia de un algoritmo sobre otro, la mayor parte de la comunidad científica se decanta por el de Kamada y Kawai. Nosotros también lo hacemos. Como muestra, un botón (figura 1).

### 3.5. Visualización

No existe una decantación clara por un formato por parte de los investigadores dedicados a la visualización de dominios. Al contrario, se utiliza una amplia variedad. Nosotros preferimos *Scalable vector graphics (SVG)* (W3C, 2004), pues se trata de un formato pequeño, veloz, y gratuito, que permite realizar ampliaciones y reducciones, así como desplazamientos en cualquier dirección, sin que por ello disminuya la calidad de los gráficos, al mismo tiempo que posibilita la búsqueda de información textual en los mismos. Además, cuenta con un firme soporte técnico y un importante apoyo por parte de la industria del sector y de muchos expertos. Sirva como referencia del auge que está cobrando las monografías de Eisenberg (2002) y Geroimenko y Chen (2005), entre otros.

## 4. Materialización y validación del cienciograma

El cienciograma de la figura 2 muestra la imagen estructural de producción científica española por medio de las categorías *ISI* durante el período 1990–2005. Su forma recuerda a una neurona humana con un gran

axón o neurita central. Dicho cienciograma, es la visualización que se obtiene como consecuencia de aplicar la metodología propuesta a los 372.002 documentos científicos españoles recogidos en las bases de datos del *ISI* entre los años 1990–2005, y agruparlos en sus respectivas categorías productivas. Los enlaces muestran las interacciones más relevantes que se producen entre categorías y expresan el punto de vista consensuado de 1.228.905 autores por medio de sus 9.550.959 citas.

Para favorecer su comprensión, cada esfera ha sido etiquetada con el nombre de su correspondiente categoría *JCR*, a la cual representa, y se le ha dado un tamaño proporcional a la cantidad de documentos que aglutina. Para ayudar a establecer visualmente la relación entre el tamaño de cada una de las categorías y su producción real, en la parte inferior izquierda de la imagen se ha representado una esfera de referencia con un tamaño equivalente a mil documentos. Las líneas que conectan las distintas esferas son las relaciones de cocitación más significativas o esenciales entre categorías, pues las más superfluas se han eliminado mediante *Pfnet*. Son más o menos gruesas dependiendo de su intensidad de cocitación (a mayor intensidad, más grosor).

La distribución espacial de las categorías en el cienciograma está determinada por el tándem compuesto por la cocitación sin normalizar de categorías y *Pfnet*. Aquellas con un mayor grado de cocitación aparecen en el centro y, en la medida en que disminuye, comienzan a distribuirse por la periferia. Al igual que **White** (2003), pero con mayores unidades de medida en nuestro caso, observamos que:

—Alrededor de las categorías más prominentes y a modo de racimos, este tándem hace aflorar las distintas áreas temáticas que componen el dominio, encadenándolas en secuencias explícitas

—El orden que las categorías tienen en ese encadenamiento no es arbitrario y revela cómo las áreas temáticas se conectan entre sí. De esta forma las conexiones entre categorías y categorías prominentes revelan áreas temáticas, y los vínculos entre categorías prominentes revelan cómo se conectan las áreas temáticas entre sí.

Por ejemplo, a simple vista y a partir de la figura 2 podemos distinguir un gran racimo central rodeado por otros menores, distribuidos por toda la superficie del cienciograma. Si nos fijamos en este racimo del centro y en otro de los que se producen en la parte central izquierda, podremos descubrir el siguiente encadenamiento: *biochemistry & molecular biology* 1 *neurosciences* 1 *psychiatry* 1 *psychology*. Este *path* o ruta, indica que en el cienciograma aparecen dos grandes

áreas temáticas que podríamos denominar como *biomedicina* y *psicología*, cuyas categorías más prominentes son *biochemistry & molecular biology* y *psychology* respectivamente, las cuales están conectadas por otras intermedias como son *neurosciences* y *psychiatry*. Lo mismo podríamos decir, entre otros, del encadenamiento que se produce en la zona inferior derecha: *mathematics applied* 1 *mathematics miscellaneous* 1 *social sciences mathematical methods* 1 *economics* 1 *history of social sciences* 1 *arts & humanities general*, que muestra cómo se conectan dos áreas temáticas que podríamos denominar como *matemáticas* y *humanidades*. Estas rutas son muy importantes, pues muestran el hilo conductor que mantiene unida la estructura científica del dominio español.

#### 4.1. Validación del cienciograma

La forma en que este cienciograma captura la estructura del dominio y distribuye la información a lo largo de ella, necesita ser contrastada. Quedan ya lejos los días en que un investigador era un experto en todos los campos de la ciencia. Otra posibilidad sería poner en común y contrastar la opinión, subjetiva, de 240 expertos distintos en la investigación nacional. Uno por cada categoría *ISI*. No obstante, la inversión económica, temporal e intelectual que esto supondría nos hace desechar esta idea. Por ello y para su validación, hemos desarrollado un método basado en un proceso estadístico como es el análisis factorial (AF):

—Realizamos AF sobre la matriz de cocitación de categorías original, es decir, sin la poda *Pfnet*.

—Extraemos el número de factores identificados.

—Etiquetamos cada uno de ellos.

—Los trasladamos al cienciograma.

El criterio adoptado para detener la extracción de factores ha sido la obtención de un *Eigenvalue* o valor propio mayor o igual a uno —aunque es simple funciona bastante bien, pues ofrece unos resultados muy acordes con las expectativas de los investigadores (**Ding; Chowdhury; Foo**, 1999)—, y se ha completado con el test *Scree* que consiste en el examen de la línea que se obtiene al representar gráficamente los valores propios de los factores identificados. La extracción de factores se detiene cuando la línea que representan comienza a nivelarse, formando prácticamente una línea paralela al eje, sin apenas pendiente (**Lewis–Beck**, 1994). Con el fin de captar la naturaleza de cada factor para poder etiquetarlo, hemos seguido el método propuesto por **Moya Anegón, Jiménez Contreras y Moneda Carrochano** (1998). Para ello hemos ordenado los factores por su peso (*factor loading*) en orden decreciente y hemos considerado como límite de pertenencia a un factor el valor igual o superior a 0,5, mien-

Factor	Nombre	Eigenvalue	% Varianza	% Varianza acumulada
1	Biomedicina	69.283	28.9	28.9
2	Ciencias de los materiales y física aplicada	30.250	12.6	41.5
3	Informática y telecomunicaciones	17.359	7.2	48.7
4	Ciencias de la tierra y del espacio	13.121	5.5	54.2
5	Psicología	10.628	4.4	58.6
6	Gestión, derecho y economía	6.975	2.9	61.5
7	Humanidades	6.740	2.8	64.3
8	Agricultura y ciencias del suelo	6.513	2.7	67.0
9	Biología animal, ecología	5.294	2.2	69.2
10	Política sanitaria, servicios médicos	3.695	1.5	70.7
11	Física nuclear y de partículas	2.865	1.2	71.9

Tabla 1. Factores del dominio científico español, 1990–2005

tras que para su denominación sólo hemos tenido en cuenta las categorías de cada factor con un valor mayor o igual a 0,7.

Para reconocer a simple vista cada una de las áreas temáticas, así como las categorías que las integran, le hemos dado el mismo color a cada una de las categorías que componen un mismo factor. Aquellas que pertenecen a más de un área se han coloreado de rojo, indicando de esta forma los puntos calientes o de interacción entre áreas temáticas a través de las categorías que comparten. Finalmente, se han puesto de color gris oscuro las que no han sido identificadas y que por tanto no pertenecen a ninguna área temática

El AF identifica muy bien en el cienciograma las áreas temáticas que podríamos denominar como consolidadas, es decir, aquellas constituidas por una categoría con un alto grado de interconexión o prominencia. Por el contrario, tiene problemas para detectar las categorías menos prominentes y, por tanto, las áreas temáticas menos consolidadas a las que dan origen, pudiendo llegar incluso a identificar como tales a especialidades dentro de una misma área temática. *Pfnet*, mediante sus racimos, muestra gráficamente las posibles áreas temáticas de un cienciograma, identifica la categoría más prominente de cada una de ellas e indica la ruta de conexión de unas áreas temáticas con otras, mediante la secuencia de categorías que las conectan. Sin embargo, no informa sobre cuáles son los límites del área temática a la que da lugar, ni tampoco de su posible denominación. Por todo ello considera-

Categorías ISI	
Acoustics	Literary reviews
Architecture	Literary theory & criticism
Art	Literature, african, australian, canadian
Asian studies	Literature, american
Astronomy & Astrophysics	Literature_german, netherlandic, scandinavian
Business	Mathematics
Computer science, theory & methods	Mathematics, miscellaneous
Demography	Optics
Engineering	Physics, fluids & plasmas
Engineering, aerospace	Physics, mathematical
Engineering, civil	Religion
Engineering, industrial	Social sciences, biomedical
Ethnic studies	Social sciences, interdisciplinary
Folklore	Thermodynamics
History & philosophy of science	Transportation
Instruments & instrumentation	Transportation science & technology

Tabla 2. Categorías no incluidas en ningún factor

mos que el AF y *Pfnet* son complementarios para la detección y análisis de la estructura de un dominio científico, pues las ventajas de uno suplen las carencias del otro, y viceversa.

En resumen, esta metodología permite validar de forma estadística (sin ningún tipo de subjetividad) la coherencia estructural del cienciograma del dominio científico español. Además, provoca la aparición de las grandes áreas temáticas que lo componen, proporcionando así una imagen de su macro-estructura intelectual subyacente, mediante lo que nosotros denominamos como cienciograma factorial (CF).

## 5. Resultados

### 5.1. Estructura de la ciencia española

El AF identifica veinticinco factores en la matriz de cocitación de 240 x 240 categorías. Mediante el test *Scree* hemos extraído once, que hemos etiquetado con el método anteriormente expuesto y que en total explican un 71,9% de la varianza (tabla 1).

El número de categorías incluidas al menos en un factor es de 208, 32 no lo han sido por ninguno (tabla 2) y 26 pertenecen a dos factores a la vez (tabla 3).

La superposición de estos resultados sobre el cienciograma original da lugar al cienciograma factorial (CF) que se muestra en la figura 3. A continuación, realizamos un paseo por su superficie. Ocupando toda la zona central y casi superior de la estructura, podemos distinguir de color púrpura al área de biomedicina. Entremezclada con ella por su parte superior y de color verde mar, encontramos política sanitaria y servicios médicos. Del mismo modo y en su zona izquierda, aparece agricultura y ciencias del suelo (verde manzana). Siguiendo el sentido de las agujas del reloj, podemos detectar fácilmente el resto de áreas temáticas. En la zona central derecha aparecen ciencias de los materiales y química aplicada (melocotón). Conectada a ésta, encontramos informática y telecomunicaciones (fucsia), física nuclear y de partículas (malva), gestión, derecho y economía (morado claro) y humanidades (azul cielo). En la zona inferior central se identifica fácilmente a biología animal y ecología (amarillo) y



Categorías / ISI	Áreas Temáticas	
Agriculture, dairy & animal science	Biomedicina	Agricultura y ciencias del suelo
Applied linguistics	Psicología	Historia y literatura
Biochemical research methods	Biomedicina	Agricultura y ciencias del suelo
Biology, miscellaneous	Biomedicina	Biología animal, ecología
Chemistry, analytical	Biomedicina	Agricultura y ciencias del suelo
Chemistry, multidisciplinary	Ciencias de los materiales, química aplicada	Agricultura y ciencias del suelo
Chemistry, organic	Ciencias de los materiales, química aplicada	Agricultura y ciencias del suelo
Communication	Ciencias de los materiales, química aplicada	Física nuclear y de partículas
Computer science, interdisciplinary applications	Biomedicina	Informática y telecomunicaciones
Ecology	Biología animal, ecología	Agricultura y ciencias del suelo
Engineering, chemical	Ciencias de los materiales, química aplicada	Agricultura y ciencias del suelo
Engineering, electrical & electronic	Ciencias de los materiales, química aplicada	Informática y telecomunicaciones
Food science & technology	Biomedicina	Agricultura y ciencias del suelo
Geochemistry & geophysics	Ciencias de los materiales, química aplicada	Agricultura y ciencias del suelo
Geosciences, interdisciplinary	Ciencias de la tierra y del espacio	Agricultura y ciencias del suelo
Mineralogy	Ciencias de la tierra y del espacio	Agricultura y ciencias del suelo
Mining & mineral processing	Ciencias de los materiales, química aplicada	Ciencias de la tierra y del espacio
Nuclear science & technology	Ciencias de los materiales, química aplicada	Física nuclear y de partículas
Operations research & management science	Gestión, derecho y economía	Informática y telecomunicaciones
Plant sciences	Biomedicina	Agricultura y ciencias del suelo
Polymer science	Ciencias de los materiales, química aplicada	Agricultura y ciencias del suelo
Psychology	Biomedicina	Psicología
Social issues	Política sanitaria, servicios médicos	Psicología
Spectroscopy	Ciencias de los materiales, química aplicada	Agricultura y ciencias del suelo
Toxicology	Biomedicina	Agricultura y ciencias del suelo
Veterinary sciences	Biomedicina	Agricultura y ciencias del suelo

Tabla3. Categorías con doble adscripción temática del dominio español 1990–2005

ciencias de la tierra y del espacio (verde grisáceo). Finalmente, en la franja izquierda, aparece situada psicología (verde esmeralda), de la que conecta una rama de las humanidades. Para establecer una rápida correspondencia entre color y el nombre de cada área temática, en la zona inferior izquierda de cada FS, hemos situado una leyenda en la que se establece su equivalencia. También se ha representado una esfera de referencia con un tamaño equivalente a mil documentos.

Teniendo en cuenta que los cienciogramas aquí representados son una esquematización extrema del dominio científico español mediante *Pfnet*, su análisis e interpretación estará basado en las inferencias que se puedan obtener de su estructura resultante. Es decir, de las categorías con un alto grado de interconexión, del principio de desigualdad del triángulo, y en las rutas de mayor peso o importancia. Es decir, si una categoría o área temática ocupan una posición central en el cienciograma, es muy probable que tanto una como otra tengan una naturaleza más general o universal en el dominio como consecuencia del número de fuentes que comparten con el resto y que por tanto, contribuyan más al desarrollo del mismo que otras que tengan una posición menos central. Del mismo modo y de forma general, cuanto más periférica sea la situación de una categoría o área temática, más exclusiva será su naturaleza, menos fuentes compartirá con el resto de categorías y menos contribuirá en el progreso del mismo.

## 5.2. Macro–estructura

Este concepto hace referencia a la forma en que las categorías *ISI* se distribuyen y agrupan, componiendo en el CF lo que nosotros denominamos como áreas te-

máticas. Lo primero que llama la atención es la combinación de pocas áreas temáticas de gran tamaño, conectadas con muchas otras más pequeñas. Se pone así de manifiesto la naturaleza hiperbólica de las distribuciones bibliométricas (**Small; Garfield, 1985**). Otro aspecto a resaltar es el patrón centro–periferia que siguen las áreas temáticas en su forma de conectarse, donde una gran área temática central hace de nodo de conexión a otras más pequeñas que la circundan. La concepción de una estructura formada por un centro y un contorno constituye un paradigma clásico y recurrente en muchos campos de la ciencia (**Borgatti; Everett, 1999**). Resulta revelador plasmar la idea clásica de la existencia de una estructura constituida por un núcleo activo, formado por un entramado denso y compacto de categorías, frente a un conglomerado disperso en sus relaciones y poco conectado.

Desde un punto de vista macro–estructural podemos distinguir tres grandes zonas. En el centro detectamos lo que podría ser denominado como ciencias médicas y de la Tierra, compuesto por: biomedicina, biología animal y ecología, ciencias de la Tierra y del espacio, psicología, agricultura y ciencias del suelo, y política sanitaria y servicios médicos. En la parte derecha podemos ver las “ciencias duras”: ciencias de los materiales y química aplicada, informática y telecomunicaciones, y física nuclear y de partículas. En la zona inferior derecha, así como en el lado izquierdo del CF, encontramos las “ciencias blandas”: gestión, derecho y economía, y humanidades. Este esquema de vertebración macro–estructural de la ciencia española coincide en gran medida con la disposición típica y persistente de los CFs de los países desarrollados, y no tanto con el de otros en vías de desarrollo, según he-

mos constatado en otros dominios (Vargas–Quesada, 2005).

La centralidad de biomedicina, puede ser interpretada como que es el área temática más universal. Es decir, se trata de aquella con la que más fuentes comparten las demás áreas y, por tanto, la que más conocimiento aporta al resto. En definitiva, la que más interviene en el desarrollo del dominio científico español. El grado de universalidad/exclusividad, del resto de áreas temáticas, estará en función de su distancia al centro. A menor distancia, mayor será su implicación en la evolución del dominio, y viceversa.

Por último, en el FS podemos encontrar también una serie de categorías de color rojo que, como dijimos anteriormente, son las responsables del intercambio de información entre áreas, indicando con su color llamativo las zonas de fricción o puntos de interacción entre las distintas áreas temáticas. Las categorías interdisciplinarias del dominio español, junto con sus correspondientes áreas a las que pertenecen, son las que se muestran en la tabla 3.

### 5.3. Micro–estructura

El concepto de micro–estructura hace referencia a la forma en que se distribuyen y componen las categorías ISI en el cienciograma. El CF de la figura 3 consta de 240 categorías y 239 enlaces que las interconectan. No aparece ninguna de ellas desconectada o aislada. Al igual que ocurría con las áreas temáticas, lo primero que llama la atención es la existencia de unas pocas categorías de gran tamaño y muchas más pequeñas. Esta mayor extensión, se produce sobre todo en el centro y centro–derecha del CF y menos en su zona izquierda e inferior derecha. Es decir, hay una mayor producción por parte de las categorías pertenecientes a las ciencias médicas y duras, que por aquellas vinculadas a las ciencias blandas. El patrón de conexión que siguen las categorías también es del tipo centro–periferia: encontramos una gran categoría central que hace de centro de interconexión a otras que están alrededor de ella, al tiempo que la convierten en responsable de mantener la cohesión estructural. A simple vista y sin lugar a dudas, la categoría más central es *biochemistry & molecular biology*. Así lo demuestra también (tabla 4) su centralidad de grado (Wasserman; Faust, 1998).

Categorías ISI	Grado
Biochemistry & molecular biology	26
Agriculture	11
Psychology	10
Medicine, general & internal	8
Physics, multidisciplinary	8
Chemistry, multidisciplinary	7
Plant sciences	7
Surgery	6
Public, environmental & occupational health	6
Geosciences, interdisciplinary	5

Tabla 4. Diez primeras categorías de mayor grado del dominio español 1990–2005

Esta máxima centralidad la convierte en la categoría más universal de las 240 que constituyen la producción científica española en el período que aquí se estudia. Se trata de la que más fuentes comparte con el resto y, en consecuencia, la que más contribuye al desarrollo científico español, de ahí su situación. Esta posición favorece su conectividad y el intercambio de conocimiento con y entre el máximo número de categorías, constituyéndose así en el eje central de la vertebración de la ciencia. El grado de participación que el resto de categorías tienen en la evolución del dominio dependerá de su distancia a esta categoría más central. Por ejemplo, *genetics & heredity*, u *oncology*, se sitúan a la distancia de un enlace; *physics atomic molecular & chemistry* están a una distancia geodésica de tres, e *information sciences & library sciences*, lo hace distanciada por once enlaces, indicándonos en cada caso su grado de participación e intervención en el desarrollo científico del dominio.

### 5.4. Columna vertebral de la ciencia española

Una de las ventajas del CF es su facilidad para mostrar la columna vertebral sobre la que se apoya la investigación del dominio que representa. Esto se consigue gracias a la habilidad que tiene *Pfnet* para seleccionar los enlaces más significativos entre categorías, junto con la posibilidad gráfica de mostrar su intensidad de cocitación, mediante el grosor de los mismos. Volviendo de nuevo al CF de la figura 3, podemos ver como hay unos enlaces más gruesos que el resto, que unen secuencias de categorías, y que insinúan la columna vertebral de este dominio. Para determinar qué enlaces y categorías son los que la constituyen, tomamos como referencia el valor más alto del enlace que une dos áreas temáticas, y eliminamos todos aquellos que quedan por debajo de dicho valor junto con las categorías que resultan desconectadas. El resultado es la columna vertebral, o la forma esencial en que se vertebra la investigación en el dominio científico español. Son tres la áreas temáticas que se distinguen (figura 4): biomedicina, agricultura y ciencias del suelo, y ciencias de los materiales y química aplicada.

El área temática que más interviene en la investigación española es biomedicina, con 10 categorías, aunque comparte tres con agricultura y ciencias del suelo, de ahí su color rojo. Obviamente, y puesto que *biochemistry & molecular biology* es la más participativa en el CF del dominio científico español completo, también lo es en su columna vertebral. La segunda área temática que más interviene en el desarrollo del dominio es agricultura y ciencias del suelo, constituida por 4 categorías que interactúan con las otras dos áreas temáticas de la columna vertebral, de ahí que todas ellas sean rojas. La tercera es ciencias de los materiales y química aplicada, formada por dos categorías

propias y una compartida (*chemistry multidisciplinary*).

## 6. Discusión

Los cienciogramas del dominio científico español ofrecen una imagen persistente de la estructura esencial del mismo, y pueden ser utilizados para ayudar a construir y/o completar la imagen mental que ya se tenía de él. Desde el punto de vista de su utilidad, son tan informativos como un mapa de metro o ferrocarril: un simple vistazo ayuda a determinar cuál es el núcleo y los límites de la línea férrea (estructura científica); resulta fácil llegar de una estación (categoría) a otra, siguiendo las vías (enlaces); es posible determinar a simple vista qué estaciones son las más importantes en función del número de conexiones que reciben y a su vez, cuántas actúan como intercambiadores (intermediarias) con otras líneas férreas, o como punto de bifurcación, según se mire.

Estos cienciogramas no son la simple representación gráfica que se obtiene como consecuencia de aplicar una serie de técnicas de reducción del espacio y de algoritmos de visualización. Son mucho más. Representan una instantánea del plebiscito u opinión consensuada que los autores de un dominio científico tienen de la imagen del mismo, expresada a través de sus citas. Por ejemplo, el hecho de que *information science & library science* esté conectada con *computer science information systems* y no con otra categoría, implica que los investigadores de este dominio consideran que, de todas las conexiones posibles que tienen ambas, en esencia, ésta es la más representativa del dominio español en el período de tiempo estudiado. Del mismo modo, el hecho de que el área temática de informática y telecomunicaciones aparezca dividida en dos: una en la parte superior derecha y otra en su zona inferior, no viene a significar más que los investigadores consideran que una parte de ella está más ligada a la ingeniería eléctrica y electrónica, mientras que otra lo está más con las matemáticas. Lo mismo podríamos decir del área de historia y literatura, mostrando así que el consenso general es que una parte de ella está más vinculada a la historia de las ciencias sociales, mientras que otra lo está más a la lingüística.

## 7. Conclusiones

En este trabajo presentamos la estructura intelectual del dominio científico español durante el período 1990–2005 mediante visualizaciones bibliométricas que hemos dado en llamar cienciogramas. Consideramos que son un buen modelo para la visualización de todo tipo de dominios, puesto que reducen el tiempo de búsqueda visual de la información, facilitan la comprensión de estructuras complejas de datos, ponen de

manifiesto las relaciones entre elementos que de otra forma no serían apreciadas, favorecen la formulación de hipótesis que pueden convertirse en objeto de análisis, debate y discusión.

La posibilidad de consultar los documentos contenidos en cada categoría, junto con los responsables de las conexiones esenciales entre las mismas (documentos contenidos en los enlaces), hace que los cienciogramas puedan ser evaluados desde el punto de vista de su significado y utilidad. Para todo aquel que lo desee, existe una versión online del cienciograma del dominio científico español, junto con la de otros dominios latinoamericanos en la web del *Grupo Scimago* (2006).

<http://www.atlasofscience.net>

**Agradecimientos:** este trabajo ha sido posible gracias a la financiación pública obtenida mediante el *Plan Nacional I+D+I: SEJ2004–08358–CO2–01/SOC*.

## 8. Bibliografía

- Borgatti, S. P.; Everett, M. G.** «Models of core/periphery structures». En: *Social networks*, 1999, v. 21, n. 4, pp. 375–395.
- Boyack, Kevin W.; Klavans, Richard; Börner, Katy.** «Mapping the backbone of science». En: *Scientometrics*, 2005, v. 64, n. 3, pp. 351–374.
- Brandenburg, F. J.; Himsolt, M.; Rohrer, C.** «An experimental comparison of force-directed and randomized graph drawing algorithms». En: *Lecture notes in computer science*, 1995, v. 1.027, n. 76, p. 87.
- Bush, Vannevar.** «As we may think». En: *The Atlantic monthly*, 1945, v. 176, n. 1, pp. 101–108.
- Börner, Katy; Chen, C.; Boyack, Kevin W.** «Visualizing knowledge domains». En: *Annual review of information science & technology*, 2003, v. 37, pp. 179–255.
- Chen, Chaomei.** «Generalised similarity analysis and pathfinder network scaling». En: *Interacting with computers*, 1998, v. 10, pp. 107–128.
- Geroimenko, Vladimir; Chen, Chaomei.** *Visualizing information using SVG and X3D: xml-based technologies for the xml-based web*. London: Springer, 2005.
- Comisión Europea (UE). Third European report on science & technology indicators, 2003: towards a knowledge-based economy*. Luxembourg: Office for Official Publications of the European Communities, 2003.
- Costa, Joan.** *La esquemática: visualizar la información*. Barcelona: Paidós, 1998.
- Crosby, A. W.** *The measure of reality: quantification and Western society*. London, Cambridge University Press, 1997.
- Ding, Y.; Chowdhury, G. G.; Foo, S.** «Mapping the intellectual structure of information retrieval studies: an author co-citation analysis, 1987–1997». En: *Journal of information science*, 1999, v. 25, n. 1, pp. 67–78.
- Eades, Peter.** «A heuristic for graph drawing». En: *Congressus numerantium*, 1984, v. 42, pp. 149–160.
- Eisemberg, J. David.** *SVG essentials*. Beijing, China: O'Reilly, 2002.
- Fruchterman, T.; Reingold, E.** «Graph drawing by force-directed placement». En: *Software practice and experience*, 1991, v. 21, pp. 1.129–1.164.
- Garfield, Eugene.** «Cienciography: mapping the tracks of science». En: *Current contents: social & behavioral sciences*, 1994, v. 7, n. 45, pp. 5–10.

## Versión online de EPI

Existe una versión electrónica de *El profesional de la información*, de uso gratuito para todos los suscriptores de la revista, que pueden acceder a través de internet a los textos completos y materiales gráficos publicados desde el año 2000.

Más información en:

<http://www.elprofesionaldelainformacion.com/contenidos.html>

**Garfield, E.; Sher, I. H.; Torpie, R. J.** *The use of citation data in writing the history of science*. Philadelphia: Institute for Scientific Information, 1964.

Grupo Scimago. *Atlas of Science*. Consultado en: 19-03-06.  
<http://www.atlasofscience.net>

**Hjørland, B.; Albrechtsen, H.** «Toward a new horizon in information science: domain analysis». En: *Journal of the American Society for Information Science (Jasis)*, 1995, v. 46, n. 6, pp. 400–425.

**Kamada, Tomihisa; Kawai, Satoru.** «An algorithm for drawing general undirected graphs». En: *Information processing letters*, 1989a, v. 31, n. 1, pp. 7–15.

**Klov Dahl, A. S.** «A note on images of social networks». En: *Social networks*, 1981, v. 3, pp. 197–214.

**Kohonen, Teuvo.** «The self-organizing map». En: *Proceedings of the IEEE*, 1985, v. 73, pp. 1.551–1.558.

**Kohonen, Teuvo.** *Self-organizing maps*. Berlin [etc.]: Springer-Verlag, 1997.

**Lewis-Beck, M. S.** *Factor analysis and related techniques*. London: Sage, 1994.

**Leydesdorff, Loet.** «Clusters and maps of science journals based on bi-connected graphs in the Journal Citation Reports». En: *Journal of documentation*, 2004a, v. 60, n. 4, pp. 371–427.

**Leydesdorff, Loet.** «Top-down decomposition of the Journal citation reports of the Social science citation index: graph and factor-analytical approaches». En: *Scientometrics*, 2004b, v. 60, n. 2, pp. 159–180.

**Leydesdorff, Loet.** «Can scientific journals be classified in terms of aggregated journal-journal citation relations using the Journal citation reports?». En: *Journal of the American Society for Information Science & Technology*, en prensa.

**Moya-Anegón, F. de; Jiménez-Contreras, E.; Moneda-Carrochano, Mercedes de la.** «Research fronts in library and information science in Spain (1985–1994)». En: *Scientometrics*, 1998, v. 42, n. 2, pp. 229–246.

**Moya-Anegón, F. de; Vargas-Quesada, Benjamín; Chinchilla-Rodríguez, Zaida; Herrero-Solana, Víctor; Corera-Álvarez, Elena; Muñoz-Fernández, Francisco J.** «A new technique for building maps of large scientific domains based on the cocitation of classes and categories». En: *Scientometrics*, 2004, v. 61, n. 1, pp. 129–145.

**Moya-Anegón, F. de; Vargas-Quesada, Benjamín; Chinchilla-Rodríguez, Zaida; Herrero-Solana, Víctor; Corera-Álvarez, Elena; Muñoz-Fernández, Francisco J.** «Domain analysis and information retrieval through the construction of heliocentric maps based on ISI-JCR category cocitation». En: *Information processing & management*, 2005, v. 41, n. 6, pp. 1.520–1.533.

**Salton, Gerard; Bergmark, D.** «A citation study of computer science literature». En: *Professional communication, IEEE transactions*, 1979, v. 22, pp. 146–158.

**Samoylenko, I; Chao, T. C.; Liu, W. C.; Chen, C. M.** «Visualizing the scientific world and its evolution». En: *Journal of The American Society for Information Science & Technology*, en prensa.

**Schvaneveldt, R. W.** *Pathfinder associative networks*. Norwood, NJ: Ablex, 1990.

**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, N. J.: Information Today, 2000.

**Small, H.; Garfield, Eugene.** «The geography of science: disciplinary and national mappings». En: *Journal of information science*, 1985, v. 11, n. 4, pp. 147–159.

The Thomson Corporation. *ISI Journal Citation Reports*. Consultado en: 14-03-06.  
<http://isiknowledge.com>

The Thomson Corporation. *ISI Web of science*. Consultado en: 14-03-06.  
<http://isiknowledge.com>

**Tufte, E. R.** *Envisioning information*. Cheshire: Graphics Press, 1994.

**Vargas-Quesada, Benjamín.** *Visualización y análisis de grandes dominios científicos mediante redes pathfinder. Tesis doctoral*. Granada: Universidad, 2005.

W3C. *Scalable Vector Graphics (SVG) 1.1 specification*. Consultado en: 14-03-06.  
<http://www.w3.org/TR/SVG11>

**Wasserman, Stanley; Faust, Katherine.** *Social network analysis: methods and applications*. Cambridge: Cambridge University Press, 1998.

**White, H. D.** «Pathfinder networks and author cocitation analysis: a re-mapping of paradigmatic information scientist». En: *Journal of the American Society for Information Science and Technology (Jasist)*, 2003, v. 54, n. 5, pp. 423–434.

**Félix de Moya-Anegón, Benjamín Vargas-Quesada, Zaida Chinchilla-Rodríguez, Elena Corera-Álvarez, Antonio González-Molina, Francisco José Muñoz-Fernández y Víctor Herrero-Solana, Grupo SCImago.**  
[felix@ugr.es](mailto:felix@ugr.es)

[benjamin.vargas@uah.es](mailto:benjamin.vargas@uah.es)

[zchinch@ugr.es](mailto:zchinch@ugr.es)

[ecorera@ugr.es](mailto:ecorera@ugr.es)

[agmo@ugr.es](mailto:agmo@ugr.es)

[franjo@ugr.es](mailto:franjo@ugr.es)

[victorhs@ugr.es](mailto:victorhs@ugr.es)

<http://www.scimago.es>

<http://www.atlasofscience.net>