

Cocitación de clases y categorías: Proyecto Atlas de la Ciencia*

Félix Moya-Anegón, Benjamín Vargas-Quesada, Zaida Chinchilla-Rodríguez,
Elena Corera-Álvarez, Francisco J. Muñoz-Fernández, Víctor Herrero-Solana**

1. Introducción

La información científica se encuentra diseminada por disciplinas que, para los no especialistas en la materia, y a veces para ellos mismos, tienen poco en común, o están poco relacionadas entre sí. Es por esto que al estudiar un dominio perteneciente a un campo de conocimiento concreto siguiendo los métodos tradicionales, en la mayoría de los casos, siempre queda la sensación de no comprender el dominio en su conjunto. Es decir, percibimos el dominio como algo desgajado de la totalidad y sin ningún tipo de relación con el resto. Es algo así como si estuviésemos completando un puzzle y al tener una pieza de él en la mano, frente a nosotros, no supiéramos el lugar donde colocarla, ni con qué otras piezas se relaciona.

La necesidad de representar la información científica de una forma más inteligible para la mente humana no es algo nuevo. *Hacer visible, especialmente para nuestra mente (aquello que no es visible para nuestros ojos)... crear una imagen mental de algo (aquello que no se ve a simple vista, una abstracción, etc.)*¹, son dos definiciones de la palabra *visualización*, que ponen de manifiesto la necesidad intrínseca de representar la información de una forma distinta a la tradicional. Pero tal y como afirma *Joan Costa*², visualizar no es el resultado implícito del acto de ver, no es un producto espontáneo del individuo que recibe la información ya visualizada. Visualizar es una tarea del comunicador visual, 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 inaprehensibles, y por tanto comprendan la información que yace oculta. El dominio de los lenguajes gráficos para visualizar estos “efectos invisibles” configuran una nueva ciencia de la comunicación visual, la esquemática, a la que *Costa* ha definido como el “tercer lenguaje” después de la imagen y el signo.

El objetivo del presente trabajo es el de proponer una nueva técnica de visualización esquemática para el análisis de grandes dominios científicos, entendidos éstos en el sentido en el que los definieron *Hjørland y Albrechtsen*³, es decir, como el reflejo de las interacciones entre los autores y su papel en la ciencia por medio de sus citas, con el

* Se desarrolla una adaptación del artículo publicado anteriormente en *Scientometrics*, v. 61, n.1, p. 129-145.

** Grupo Scimago (<http://scimago.ugr.es>, <http://www.atlasofscience.net>), Departamento de Biblioteconomía y Documentación, Universidad de Granada (España).

¹ *Framework of a visualization system*, disponible electrónicamente en http://www.siggraph.org/education/materials/HyperVis/abs_con1/main.htm.

² Thomson ISI, *ISI Web of Science* [Web page], Available at: <http://wos.cica.es/> [Consultado en: 2003].

³ E. Garfield, I.H. Sher, y R.J. Torpie, *The use of citation data in writing the history of science*, Philadelphia, Institute for Scientific Information, 1964.

objeto de construir el atlas de la ciencia, poder realizar su análisis, y que eventualmente sea utilizado como intermediario para la recuperación de información.

Comenzaremos haciendo una breve revisión bibliográfica de los diversos intentos y aportaciones realizados en aras de la representación de grandes dominios científicos y de cómo éstos, mientras las herramientas y técnicas alcanzan su grado de madurez, se han centrado en la visualización de pequeños dominios. Continuaremos con una propuesta sobre qué elementos y unidades de medida utilizar para la representación gráfica de grandes dominios. Seguiremos con la metodología desarrollada para la generación de mapas que favorecen e invitan al *browsing*. Finalizaremos con los resultados y conclusiones, mostrando varios ejemplos de dichos mapas, comentando sus características y peculiaridades.

2. Breve revisión bibliográfica de los intentos y aportaciones en la representación de grandes dominios científicos

El primero en articular esta necesidad fue *Doyle* en 1961⁴, que resalta la importancia de los ordenadores para producir mapas similares a los que genera el cerebro, indica cómo estos pueden ser proyectados en espacios multidimensionales y da su opinión sobre el objetivo de dichos mapas y cómo construirlos. Realizar la visualización de un gran dominio científico, o *the big picture* como también se le ha denominado, ha sido un anhelo perseguido durante mucho tiempo. *Garfield*[5], en un artículo publicado en *American Documentation*, mostraba su gran interés por la construcción de mapas históricos mediante el uso de citas. Interés que, junto a *Irv Sher* en 1964, se plasmó en la realización de un mapa histórico en el que se mostraba el desarrollo del ADN desde *Mendel* hasta *Niremberg*. Ese mismo año, *Garfield* junto con *Sher* y *Torpie*[3], generan manualmente distintos mapas histórico topológicos de la ciencia, a partir de las citas de la producción científica sobre la investigación en el ADN, utilizando como variable el emparejamiento bibliográfico –bibliographic coupling-. *DeSolla Price*[6], demostró que los patrones de citas utilizados por los autores de artículos científicos, definían los frentes de investigación y que éstos podían ser aprovechados para delinear una topología que reflejase la estructura de la producción científica. Pero el verdadero avance para el diseño de mapas o gráficos de un dominio, desde nuestro punto de vista, vino de la mano de *Small*[7] y *Marshakova*[8], al proponer de forma independiente la cocitación de documentos como variable de estudio en los análisis de citas de la producción científica. Los mapas de la ciencia, en donde se muestran todas las especialidades de las ciencias naturales, elaborados por *Small* y *Griffith*[9] o *Griffith*, *Small*, *Stonehill* y *Dey*[10], utilizando como fuente de información el *Science Citation Index (SCI)* y la cocitación como variable de relación, se convierten en un hito fundamental en el desarrollo de la representación de dominios científicos. Lo verdaderamente importante de la metodología utilizada por los autores anteriormente citados, es que identificaba perfectamente los grupos de documentos que tenían intereses intelectuales comunes, mostrando la prueba de que la ciencia es una red de especialidades interconectadas entre sí, que se pueden contemplar utilizando como base el análisis cuantitativo de la producción escrita. En 1975, haciéndose eco de los trabajos de los autores anteriormente citados, *Aaronson*[11] radiografía la biomedicina

⁴ L. B. Doyle, Semantic roadmaps for literature searchers, *Journal of the Association for Computing Machinery*, 8, (1961) No. 4, 553-578.

de los años 1972 y 1973, pudiéndose apreciar su evolución en el tiempo y observar, en el mapa de 1973, lo que el autor llamó el *supercluster*, al mostrar la convergencia de otras especialidades. Los mapas de *Aaronson* son también un hito en la historia de la representación gráfica de un dominio, no sólo por reflejar la evolución del mismo a través del tiempo, o por mostrar como las distintas disciplinas interactúan entre sí, también lo son por la información que suministran, pues los documentos aparecen agrupados conformando un mapa de *clusters*, en donde cada uno de ellos tiene su propia denominación y número de documentos que lo componen, así como el grado de conexión entre ellos, representado por líneas acompañadas del valor de cocitación que los conectan. Además, permiten detectar los *cluster* más importantes, pues son aquellos que tienen un mayor número de relaciones con el resto. Poco después de la aparición de los mapas de *Aaronson*, *Garfield*[12], informa de que el *ISI (Institute for Science Information)*, está trabajando en la elaboración de un *Atlas de la Ciencia*. Pero hubo que esperar seis años para que este deseo se hiciera realidad. Fue en 1981[13] cuando aparecieron los dos primeros volúmenes del *Atlas*, en concreto los referentes a la bioquímica y a la biología molecular. Las técnicas utilizadas para la generación de los mapas del *Atlas* no difieren mucho de las utilizadas por *Small* y *Griffith* en 1974. Es decir, generación de *clusters* a partir de los datos obtenidos por la cocitación de documentos en una disciplina concreta, aunque incorporan la novedad de un nuevo posicionamiento espacial de los *clusters*, mediante el uso de técnicas de escalamiento multidimensional (*MDS*). Con el paso de los años han ido apareciendo distintos volúmenes del *Atlas de la Ciencia*: Biotecnología y genética molecular[14], Bioquímica, inmunología y biología animal y vegetal[15], quedando paralizada la producción del mismo, hasta la fecha, con el volumen referente a Farmacología[16]. Mientras tanto, *Small*, ha seguido trabajando en el diseño de mapas de dominios científicos, depurando y mejorando las técnicas utilizadas en sus primeros mapas, convirtiéndose en el principal especialista del *ISI* en la investigación y elaboración de mapas de la ciencia. Los siguientes trabajos no son más que una breve muestra de su magna producción: [17-24].

A partir de los años 90, con la aparición de nuevos métodos de recopilación de información, así como de nuevas técnicas para el análisis, visualización y posicionamiento espacial, de la información, que *Börner*, *Chen* y *Boyack*[25] revisan de forma magistral, comienzan a proliferar los estudios basados en técnicas de visualización de la estructura de pequeños dominios científicos, que a la vez permitan la clasificación y/o recuperación de la información que contienen, dejando a un lado, o aplazando hasta que las herramientas sean lo suficientemente consistentes, la elaboración de mapas de grandes dominios científicos o *The big picture*. Así, por ejemplo, y mencionando sólo a algunos, *Braam*, *Moed* y *Raan*[26; 27], proponen el uso combinado de la cocitación con el análisis de palabras –co-word análisis–, para la generación de mapas de la ciencia, haciendo hincapié en su estructura y aspectos dinámicos. *Lin*, *Soergel* y *Marchionini*[28], desarrollan un mapa autoorganizativo (*SOM*) que representa de forma gráfica y esquemática las relaciones semánticas entre los documentos, pudiendo ser utilizado como interfaz bibliográfica para la recuperación online de la información que representa. *Hjorland* y *Albrechtsen*[29], plantean un nuevo modelo o enfoque para el estudio de la documentación, basado en el estudio de las áreas de conocimiento, desde el punto de vista de la forma de pensar o hablar que tiene la sociedad, que está determinada por su entorno laboral, social, económico, etcétera, y que ellos denominan análisis de dominios. *White* y *McCain*[30], utilizando como base el análisis de dominios de los autores anteriormente citados, proponen la representación gráfica o visualización, como modelo para el análisis y recuperación de la información que el usuario estime oportuna. Para corroborar su teoría, en 1988 realizan el análisis de un dominio concreto: la documentación[31], utilizando como unidad de análisis los autores de las 12 revistas que, según su criterio, fueron consideradas más importantes en el período de tiempo comprendido entre 1972 y

1995. *Garfield*[32] dice que las nuevas técnicas de visualización hacen posible la generación de mapas globales de la ciencia que, al hacer zoom sobre ellos o representar distintos períodos de tiempo, permiten identificar los frentes de investigación emergentes, que no son más que el reflejo de las preocupaciones de los investigadores en sus trabajos científicos, por lo que es posible asociar nombres de autores a cada frente. *White, Lin, y McCain*[33], comparan el modo tradicional de visualizar los dominios científicos: (MDS), con los mapas auto-organizativos (SOM), llegando a la conclusión de que los resultados son muy similares, pero con la diferencia de que el último permite una más fácil integración entre información bibliográfica y recuperación de la misma. Para *Ding, Chowdhury, Foo y Qiang*[34], el uso de las técnicas bibliométricas permite la simplificación de un área de conocimiento a sus elementos principales, la representación gráfica de dichas áreas y de sus correspondientes subáreas, así como su utilización para una mejor comprensión del dominio por parte del usuario. *Merton*[35], dice que lo que en un principio sólo fue concebido por *Garfield*, para recuperar información, es en realidad una magnífica herramienta para el estudio de la sociología de la ciencia. *White*[36], da a conocer las redes centradas en un sujeto, argumentando que, hasta ahora, los mapas de un dominio siempre se habían hecho a partir de los autores más citados de un grupo seleccionado de publicaciones, proponiendo como alternativa que éstos se hagan a partir del nombre de un autor, que suministraría el usuario. Consiguiéndose así una menor carga cognitiva por parte del que realiza la consulta, el desarrollo de interfaces para usuarios inexpertos que faciliten la recuperación de la información a partir de información bibliográfica, así como la posibilidad de que dichos interfaces se puedan generar de forma dinámica. Para *Noyons, Moed y Luwel*[37], *Buter y Noyons*[38], y *Noyons*[39], el uso de los mapas como metáforas de una disciplina científica tienen un gran potencial como interfaz para la recuperación de la información, no obstante, éstos tienen algunas limitaciones como consecuencia de que el usuario no comprende bien lo que está viendo, por lo que aportan algunas soluciones que permiten una mejor exploración del dominio. *Chen y Paul*[40] describen como, para hacer mapas del conocimiento claros y fáciles de interpretar, *han desarrollado un método que amplía y transforma el análisis tradicional de coautores en patrones estructurales de la literatura científica que pueden ser representados en mapas 3D*. Para *Chen, Paul y O'keefe*[41], la proliferación de las técnicas de visualización de la información, han permitido representar la esencia del conocimiento, jugando un papel fundamental en el proceso de modelización y representación de la estructura o mapa intelectual de un determinado dominio, ya sea geográfico, temático o intelectual. De nuevo *Ding, Chowdhury y Foo*[42], hacen un mapa de la estructura intelectual del campo de la recuperación de la información, en un período de 10 años (1987-1997), mostrando modelos, patrones y tendencias de dicho campo, así como distintas medidas del grado de asociación entre los términos más relevantes de los documentos producidos bajo el epígrafe "recuperación de información". *Ingwesen y Larsen*[43], investigan las ventajas de la visualización gráfica de la producción científica de un área geográfica, con respecto a los métodos tradicionales. Para ello, utilizando MDS, hacen un mapa de la producción científica de diecisiete países europeos, en nueve áreas de las ciencias sociales. *Guerrero Bote, Moya Anegón y Herrero Solana*[44], describen un método en el que los documentos de una base de datos pueden ser clasificados de forma automática mediante un SOM y de cómo éste puede ser utilizado para el *browsing* y la recuperación de información en esa base de datos. *White, Buzydlowski y Lin*[45-47], basándose en la experiencia acumulada por *White* en los trabajos anteriormente citados y en sus *CAMEOs*[48], implementan un sistema dinámico de visualización: *Authorlink*, basado en la cocitación de autores, que permite el *browsing* y la recuperación de información en una base de datos constituida por los registros de diez años del *Arts & Humanities Citation Index (A&HCI)*, en tiempo real. *Small*[49], teoriza sobre el diseño de una herramienta web capaz de detectar y monitorizar en tiempo real, los cambios que se producen en los frentes de investigación, como

consecuencia de sus interacciones. *White*[50], propone Pathfinder Networks (PFNET) como una nueva técnica para realizar mapas de cocitación de autores (ACA), detallando sus ventajas. *Chen y Kuljis*[51], mediante el uso de la citación y la cocitación en el campo de la física, estudian la aparición y evolución de nuevos frentes de investigación a través del tiempo. *Morris, Yen, Wu y Asnake*[52] trabajan en la visualización, detección e identificación de los cambios temporales en los frentes de investigación. *Boyack y Börner*[53], con fines evaluativos, generan mapas de publicaciones científicas subvencionadas, haciendo visible la relación entre financiación gubernamental y número de citas recibidas. En resumen, los mapas o visualizaciones de dominios se están utilizando en la actualidad para mostrar relaciones entre documentos, detectar los autores más importantes de una determinada disciplina, o analizar la estructura de un área de conocimiento, y su evolución mediante su representación en sucesivos espacios temporales. Las técnicas que se están utilizando son *clustering*, *MDS*, análisis de factores, o redes sociales basadas en modelos de grafos, cada una de ellas por separado o combinadas de forma indistinta.

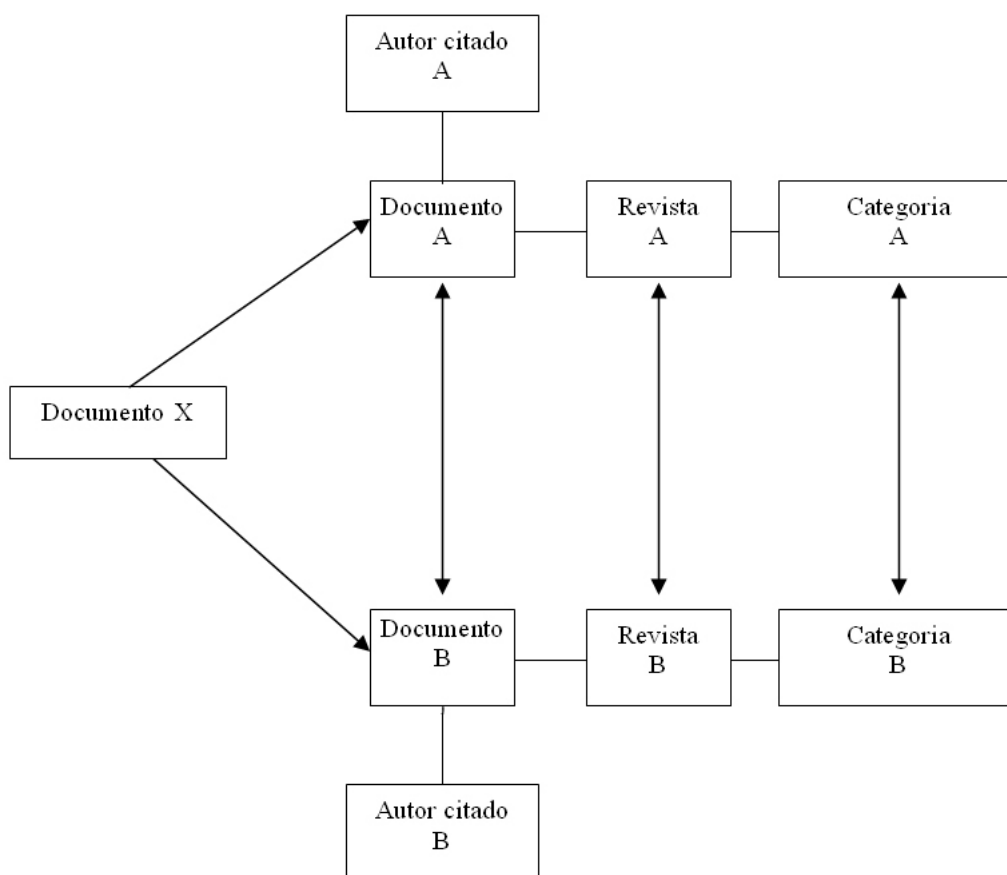
El objetivo fundamental que se persigue hoy día, es el de hacer una primera representación gráfica o mapa de un dominio que sea lo suficientemente general e informativo para que el usuario tome conciencia del área en la que trabaja, y luego hacer zoom o ir descendiendo niveles en cada una de las disciplinas resultantes, ya sea mediante el uso de las técnicas multivariante, o por medio de trazadores de redes.

3. Nuestra propuesta

Desde nuestro punto de vista, hoy por hoy, la cocitación es la mejor herramienta con la que contamos para obtener la información relacional de los documentos, a partir de la cual representar de forma esquemática la imagen de un dominio. Así, en función del tipo de variable analizada: palabras, documentos, autores o revistas, es posible mostrar la estructura intelectual subyacente de un dominio, por medio de las relaciones existentes entre los elementos que lo componen. El uso de una variable u otra en la cocitación, depende en gran medida del tamaño del dominio a representar, así como de la posibilidad o interés por conseguir visualizaciones dinámicas o no, es decir, obtener mapas *online* u *offline*[30]. Pero también hay que tener muy en cuenta que nos seguimos encontrado con la limitación física, a la que *Tufte*[54; 55] lleva dedicado más de una década, que es la representación de la información en un espacio reducido de poca resolución.

Nuestra propuesta consiste en la representación gráfica y esquemática de grandes dominios temáticos, geográficos, institucionales, etcétera, basada en la cocitación, para construir el atlas de la ciencia. Resulta obvio que es necesario conseguir algún tipo de aglomeración o clusterización, si lo que se pretende es representar en la pantalla de un ordenador la estructura intelectual de un gran dominio y que a la vez sea inteligible para la mente humana. Por ello, proponemos la cocitación de clases y de categorías, que se basa en el esquema representado en el cuadro 1.

Cuadro 1



Es un hecho aceptado por la comunidad investigadora, en el área de la visualización y representación de la información, que la frecuencia con la que dos documentos cualesquiera son citados conjuntamente, representa el grado de afinidad de los mismos según el punto de vista del autor/es citantes, que mide su grado de relación o asociación, tal y como es percibida por dicho autor/es y que representa de forma esquemática la estructura del objeto de estudio.

En el cuadro 1, los documentos A y B son cocitados por el documento X. La fuerza o intensidad con la que A y B se relacionan, depende del número de veces que A y B son cocitados. De la misma forma que la comunidad investigadora acepta que el documento A y el documento B, son una entidad de cocitación válida para representar la estructura de un dominio a la vez que pone de manifiesto sus relaciones semánticas e intelectuales, utilizando los documentos como unidad de medida, también acepta el uso de los autores como entidad de cocitación y como unidad de medida para representar dicho dominio. Lo mismo ocurre con el caso de las revistas.

El *ISI*, en el *JCR (Journal Citation Report)*[56], incluye cada revista en categorías temáticas –subject categories- en función de las áreas temáticas o disciplinas que abarcan los contenidos de los artículos que se publican en ellas. Por ejemplo, la revista *Journal of the American Society for information Science and Technology*, está adscrita por el *ISI-JCR*[56] a las categorías: *Computer Science and Information System e Information Science and Library Science*. Estableciendo un paralelismo con la casuística anterior, podríamos decir que la cocitación de categorías asignadas por el *ISI-JCR*, puede ser utilizada como entidad de cocitación siguiendo el orden establecido en el cuadro 1, es decir, documentos, autores, revistas, materias *ISI-JCR*, y que también puede ser utilizada como unidad de medida válida para representar la

estructura intelectual de un dominio. Retomando el ejemplo anterior, si los documentos A y B, que aparecen publicados en la revista *Journal of the American Society for Information Science and Technology*, son cocitados por el documento X, está claro que existe una relación entre las categorías *Computer Science and Information System* e *Information Science and Library Science*, con una intensidad que estará en función del número de veces que ambas categorías sean cocitadas. Pero, si el documento A, ha sido publicado en la revista anteriormente citada y por tanto tiene asignadas las categorías ya mencionadas, y el documento B se ha publicado en la revista *European Journal of Operational Research*, que tiene asignadas por el *ISI-JCR* las materias *Management* y *Operations research & Management Science* se nos indicará así, que existe una relación entre las categorías *Computer Science and Information System*, *Information Science and Library Science*, *Management* y *Operations research & Management Science*, cuya fuerza o intensidad dependerá del número de veces en que dichas categorías sean cocitadas, o lo que es lo mismo, del número de veces en que documentos publicados en revistas pertenecientes a esas categorías sean cocitados.

Existen diferentes clasificaciones que agrupan las categorías temáticas del *ISI-JCR* en conglomerados superiores o áreas temática –*topic areas*-. Nosotros, puesto que pretendemos representar la estructura de la producción científica española, hemos adoptado la clasificación de la *ANEP* (Agencia Nacional de Evaluación y Prospectiva), que adopta las categorías *ISI-JCR* como elemento distintivo y aglutinador de la producción científica en España, agrupándolas en veinticinco grandes área temáticas o clases –*Topic Areas*-. Al igual que ocurre en el *ISI-JCR*, una misma categoría puede pertenecer a distintas áreas temáticas. Siguiendo el razonamiento anterior utilizado en las categorías, las clases pueden ser perfectamente utilizadas como entidad de cocitación y como unidad de medida mayor, válida para representar la estructura intelectual de un dominio, como demostraremos más adelante.

En definitiva, proponemos el uso de la cocitación de clases y de categorías como entidades de cocitación y unidades de medida, para la generación de grandes visualizaciones esquemáticas que actúen como interfaces para el análisis de dominios.

4. Metodología

Con propósitos estrictamente académicos y de investigación, hemos descargado del *Web of Science*[2], en concreto del *SCI-E* (*Science Citation Index-Expanded*), *SSCI* (*Social Science Citation Index*) y *A&HCI* (*Arts & Humanities Citation Index*), los registros de la producción científica española en el año 2000, y los hemos introducido en una base de datos construida *ad hoc*, para realizar las consultas necesarias. La base de datos cuenta con 172.562 autores, que han publicado 26.062 documentos (artículos, biographical-items, book reviews, corrections, editorial materials, letters, meeting abstracts, news items y reviews) recogidos en 3.838 revistas *ISI*, a las que al asignarle las 243 categorías contempladas en el *ISI-JCR* del año 2000, han quedando reducidas a 222 categorías, al no contar España en el 2000, con producción científica en aquellas revistas *ISI* a las que corresponden las 21 categorías restantes. Las 222 categorías resultantes han sido agrupadas, según la clasificación de la *ANEP*, en 25 grandes clases temáticas, teniendo en cuenta que una misma categoría puede pertenecer a distintas clases.

Puesto que pretendemos poner de manifiesto las relaciones existentes entre las diferentes disciplinas recogidas en el *SCI*, *SSCI* y *A&HCI*, es decir, entre las ciencias naturales, las ciencias sociales y el arte y las humanidades, lo primero que debemos de hacer es resolver el problema de la normalización de los índices de citación entre disciplinas, tal y como sugieren *Small* y *Garfield* en su trabajo: *The geography of science: disciplinary and national mappings*[18]. Para ello, al realizar la consulta de cocitación correspondiente de clases o de categorías, dentro de la propia consulta, normalizamos esta medida de asociación al dividir la cocitación por la raíz cuadrada del producto de la frecuencia de las citas de los documentos cocitados.

$$C_c = \frac{C_{cij}}{\sqrt{c_i \cdot c_j}}$$

Medida de Cocitación normalizada[12].

Donde:

- C_c es Cocitación
- c citación.

La generación de los mapas para la visualización de los dominios la hacemos mediante el uso de redes sociales. Para ello utilizamos *Pajek*[57], que es un programa para la visualización y análisis de grandes redes sociales y gratuito para usos no comerciales, que genera redes sociales desde un punto de vista estético: utilizando el máximo de espacio disponible, minimizando el número de enlaces cruzados y forzando la separación de los nodos. Por indicación nuestra, utiliza el algoritmo de *Kamada Kawai*[58], para la generación automática de representaciones en un plano, partiendo desde una posición circular de los nodos.

4.1. Mapa de Cocitación de Clases (First Level)

Para obtener los datos básicos necesarios con los que representar de forma gráfica el dominio de España, realizamos una consulta de cocitación de clases o áreas temáticas, según la clasificación de la *ANEP*, siguiendo el criterio antes mencionado.

El resultado es una matriz de cocitación simétrica de 24 por 24 clases o áreas temáticas. Ciertamente, el número de coincidencias que muestra la matriz de cocitación entre las clases resulta elevado, dificultando, en algunas zonas, la visualización clara de la estructura del dominio, debido al fuerte intercambio intelectual que existe entre algunas clases. Para aclarar la visualización y siguiendo el consejo de *Small*[24], *...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*, hacemos una poda de las relaciones entre las clases utilizando *Pathfinder Network (PFNET)*. Pero, tal y como argumenta *White*[50], *este algoritmo poda todos los paths, excepto aquellos que tienen el grado de cocitación más alto*, por lo que el número de relaciones o coincidencias que quedarían en la matriz es muy reducido. Puesto que pretendemos representar algo más que la estructura básica de un dominio, hemos optado por utilizar *PFNET* con valores de poda menos drásticos, es decir, $q = n - 2$, para $r = \infty$. De esta forma, conseguimos mostrar un número más elevado de enlaces, sin dificultar la visualización clara del dominio.

Pajek hace cada nodo del mismo color y tamaño, por lo que a cada clase se le asigna un color para distinguirla del resto, un tamaño proporcional a la totalidad de la producción científica de España en el año 2000, así como el nombre asociado a la clasificación de la *ANEP*. Esta información, junto con la matriz de cocitación, es procesada por dicho programa. El resultado final es una red social, en la que cada clase está representada por un nodo que se conecta con otros nodos o clases a través de enlaces no dirigidos, indicando las relaciones que existen entre ellas, así como su intensidad, mediante el grosor de los mismos. Sólo en un par de casos en los que la etiqueta de un nodo aparece parcialmente superpuesta con otra, se ha recurrido a modificar levemente, de forma manual, la ubicación asignada por el algoritmo de *Kamada Kawai*. La red definitiva es exportada a un formato *SVG (Scalable Vector Graphics)*[59], que nos permite hacer zoom y desplazamientos del gráfico en cualquier dirección dentro de la pantalla.

4.2. Mapa de Cocitación de Categorías (Secondary Level)

El proceso a seguir es muy similar al que se utiliza en las clases, sólo difiere en pequeños detalles destinados a mejorar la visualización de las relaciones intelectuales. Como hemos comentado anteriormente, cada una de las clases o áreas temáticas están compuestas por un número determinado de categorías *ISI*, asignadas según el criterio de la *ANEP*. Para cada clase *ANEP*, hacemos una consulta de cocitación de categorías *ISI*, normalizada según los criterios antes expuestos, con lo que obtenemos una matriz de cocitación simétrica de $N \times N$ categorías, en función del número de las mismas que aglutina cada clase. Realizamos la poda mediante *MST*. Asignamos color a cada categoría, que es igual al de la clase a la que pertenece. Indicamos el tamaño, pero con una pequeña salvedad con respecto al caso anterior: dentro de cada clase, a la categoría que tiene una mayor producción científica, se le asigna el mayor tamaño y el resto de las categorías se hace proporcional a ésta, dependiendo de la producción científica de cada una de ellas. De esta forma, como no suelen existir grandes diferencias de producción entre las categorías de una misma clase, podemos conseguir que categorías con una producción científica pequeña, por ejemplo 4, sean perfectamente visibles.

Suministramos a *Pajek* el nombre de las categorías que componen cada clase, su tamaño, color y la matriz de cocitación correspondiente, que es la que establece las relaciones entre categorías. En este caso, se indica al software que dichas relaciones sean dirigidas. Aunque se trata de una matriz simétrica y por tanto cuando existan relaciones entre categorías, éstas también lo serán, esta pequeña licencia nos permite, como veremos más adelante, obtener unos resultados muy buenos en la claridad y estructura de la visualización del dominio. No obstante, el distinto grosor de los enlaces sigue indicando el grado de intensidad entre las relaciones.

La red social que obtenemos, modificada sólo en algunos casos para evitar la superposición de etiquetas de los nodos o categorías, es exportada a formato *SVG*.

4.3. Mapa de Vecinos (Third Level)

Para este caso no tenemos en cuenta la clasificación de la *ANEP*, sólo utilizamos las categorías *ISI*. Pues nuestro objetivo es mostrar qué categorías son las que más se relacionan entre sí, en qué medida, y qué documentos son los que se “ocultan” detrás de cada categoría y de cada uno de los enlaces o relaciones. Para ello, creamos una

red egocéntrica o en forma de estrella, de cada una de las categorías que tienen producción en España en el año 2000.

Partimos de una matriz de cocitación de 222 x 222 categorías. A partir de ella, construimos una lista de vecinos basada en un sujeto, que en cada caso será el de la categoría objeto de estudio o representación. Procesamos con *Pajek* la lista de vecinos, el nombre de todas las categorías, color y tamaño, estos dos últimos establecidos de la misma forma que en el caso anterior. Puesto que sería imposible mostrar de una forma clara e inteligible las 222 categorías, realizamos un proceso de poda. Pero no con *MST*. Para este caso utilizamos como umbral el valor obtenido de calcular la media más la desviación estándar de aquellas categorías que, dentro de la matriz de cocitación completa, pertenecen al vector de la categoría que vamos a representar, y que tienen un valor mayor que cero. Es decir, utilizamos un mecanismo de poda con un umbral variable que, es establecido por la propia idiosincrasia de cada categoría central, con respecto al resto. Eliminamos los enlaces y vértices que se conectan a la categoría o nodo central, con un umbral inferior al mencionado. Para obtener la representación de la red, como en los casos anteriores, recurrimos al algoritmo de *Kamada Kawai*, pero en esta ocasión le indicamos que el valor de los enlaces no es cocitación, sino similitud. De esta forma, el grosor de las relaciones es siempre el mismo, pero su longitud varía. Pudiendo detectarse a simple vista qué categorías son las que están más cercanas a la central, y por tanto tienen una mayor intensidad o relación. En algunos casos, en los que enlaces o etiquetas de nodos aparecían superpuestos, se han modificado levemente de forma manual. El mapa resultante es exportado a formato SVG.

5. Resultados

Pensamos que este tipo de representaciones permiten, a cualquier persona que se sitúe frente a ellas, detectar de forma clara y rápida, la estructura de un dominio por medio de sus nodos, enlaces e intensidad de los mismos.

5.1. Mapa de Primer Nivel o Cocitación de Clases

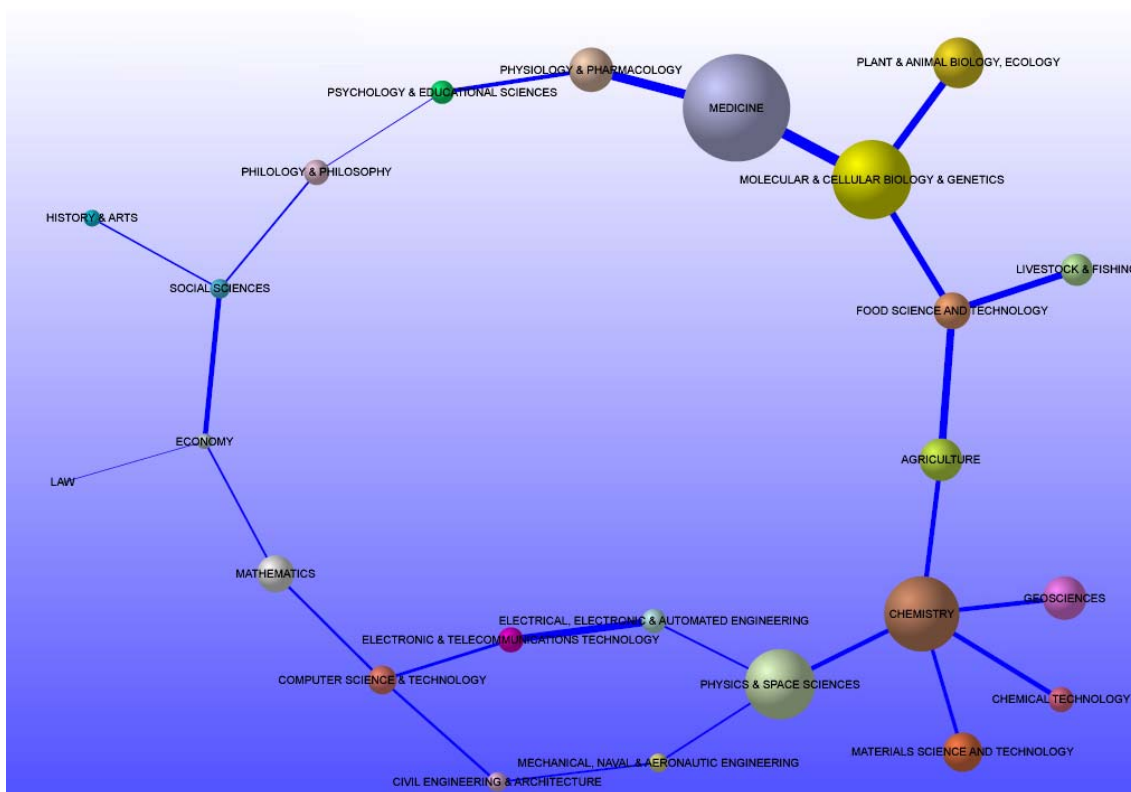


Figura 1. Mapa de primer nivel o de cocitación de clases (España, 2000)

Cada clase, está representada por un nodo en la red con su nombre correspondiente. A simple vista, se puede observar cómo la producción científica española en el año 2000, aparece distribuida en todo lo ancho y largo del gráfico, ocupando el máximo de espacio disponible. Podemos ver como en el centro aparece un gran “agujero”, indicando que no existe ninguna clase o área temática que esté relacionada de un forma directa con todas, al igual que ocurre con las representaciones en *MDS*. Ni siquiera *Chemistry*, que es la clase más relacionada, ocupa esta posición central. Pues sus conexiones con el resto la obligan a adoptar una posición más periférica.

Desde un punto de vista general, podemos ver claramente reflejadas las tres bases de datos del *ISI*. En la zona superior izquierda aparece claramente delimitada la base de datos del *SSCI* y dependiendo de ella el contenido de *A&HCI*, mientras que el resto lo ocupan las áreas temáticas que integran las categorías de la base de datos *SCI*. Dentro de lo que podemos denominar zona *SCI-E*, distinguimos tres grandes bloques: por un lado está al que llamaremos Ciencias de la vida que está compuesto por: *Agricultura*, *Livestock and Fishing*, *Food Sciences and Technology*, *Molecular & Cellular Biology & Genetics*, *Plant & Animal Biology & Ecology*, *Medicine*, *Physiology and Pharmacology* y *Psychology and Educational Sciences*; otro bloque es el de la Física, Química y Ciencias de la Tierra y del Espacio que se compondría por: *Chemistry*, *Geosciences*, *Chemical Technology*, *Materials Science & Technology* y *Physics and Space Sciences*; por último está el grupo de las ingenierías y la informática: *Electric, Electronic & Automated Engineering*, *Electronic & Telecommunications Technology*, *Computer Sciences & Technology*, *Civil Engineering & Architecture*, *Mechanical Naval & Aeronautical Engineering*, y *Mathematics*. Resulta curioso observar como la red, y en definitiva las tres bases de datos del *ISI*, están conectadas por nodos puentes, o puntos de corte: *Psychology and Educational Sciences* y *Mathematics*, que conectan los tres grandes grupos antes mencionados,

de tal forma que si eliminásemos sus enlaces con el resto, dejarían la red descompuesta en tres componentes, que coinciden con las tres bases de datos del *ISI*.

Los enlaces que unen los distintos nodos, indican las relaciones que existen entre ellos. La intensidad de las relaciones está determinada por el grosor de los enlaces. Por ejemplo, a simple vista podemos observar como el enlace que existe entre *Social Sciences* y *Economy*, es más grueso que el que tiene *Social Sciences* con *History & Arts*, indicando de forma intuitiva que la relación entre esos nodos es más intensa que las que tiene con el resto. Si se mira en la parte inferior del gráfico, veremos como ocurre lo mismo con *Electronic & Telecommunications Technology*, *Electric*, *Electronic & Automated Engineering* y *Computer Sciences & Technology*. De esta forma, a la vez que se aprecia que entre ellos mismos existen enlaces de gran intensidad, se pone de manifiesto que el número de fuentes que comparten, en este caso categorías, es más elevado que con el resto.

5.2. Mapas de Segundo Nivel o Cocitación de Categorías

Cada uno de los nodos del mapa de cocitación de clases, da lugar a la generación de un nuevo mapa de cocitación de categorías *ISI*, que hacen un total de 25 mapas, tal y como explicamos en el apartado de la metodología.

Como ejemplo y por afinidad al área en la que investigamos, mostraremos el mapa de las Ciencias Sociales, que incluye la categoría *Library & Information Sciences* (Figura 2).

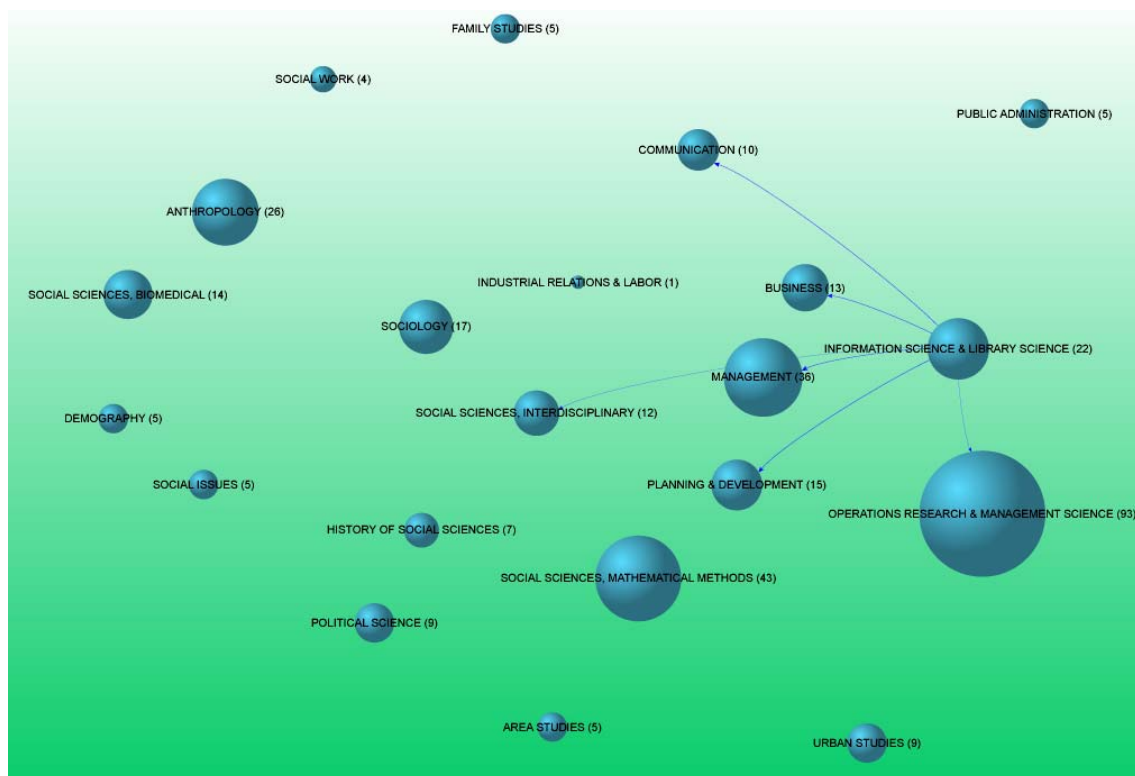


Figura 2. Mapa de Cocitación de Categorías. Social Sciences

Cada nodo de la red representa una categoría, cuyo nombre ha sido asignado por el *ISI-JCR* y junto a él, hemos añadido la producción científica española de 2000 en cada una de las mismas. Como es típico en este tipo de redes, la representación ocupa al máximo el espacio disponible, al mismo tiempo que los nodos con un mayor número de enlaces tienden a ocupar posiciones centrales, mientras que los que tienen un menor número, se sitúan en la periferia. En el caso de las Ciencias Sociales, los nodos que aparecen en la zona central son, por este orden: *Sociology*, *Planning & Development*, *Social Sciences Mathematical Methods*, *Social Sciences-Interdisciplinary* y *Management*. El resto, incluida *Library & Information Sciences*, ocupan posiciones más o menos periféricas.

Al igual que ocurría en el mapa de cocitación de clases, la intensidad de las relaciones está determinada por el grosor de los enlaces. Pero a diferencia de él, y sólo para los mapas de cocitación de categorías, hemos incluido una pequeña aplicación que nos permite ocultar los enlaces con el fin de mejorar la visibilidad, y que éstos sólo sean visibles cuando el cursor del ratón se posiciona sobre alguno de los nodos o etiquetas, tal y como se puede observar en la figura 2 y 2.1.

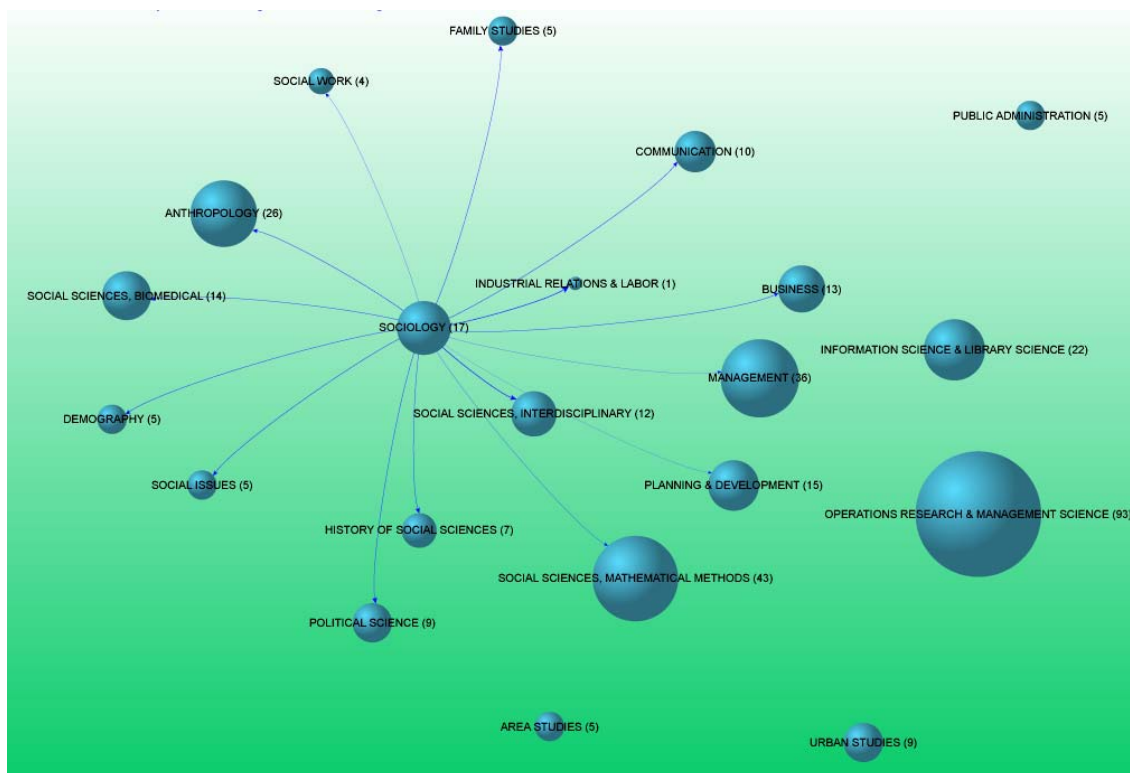


Figura 2.1 Mapa de Cocitación de Categorías. Social Sciences

Para el caso de *Library & Information Sciences*, de los seis enlaces que mantiene con otros nodos, dos son bastante débiles: *Operation Research & Management Science* y *Social Sciences Interdisciplinary*, mientras que los otros cuatro, aunque también débiles, son más intensos que los dos primeros. En el caso de *Sociology*, de las quince relaciones que mantiene, dos son especialmente fuertes: *Social Sciences Interdisciplinary* e *Industrial Relations & Labor*.

5.3. Mapas de Vecinos

Cada uno de los nodos de los mapas de cocitación de categorías, se convierten en el punto de partida de un nuevo mapa, que nosotros llamamos mapas de vecinos, dando lugar a 222 mapas en total.

Siguiendo con el área de *Library and Information Sciences*, mostramos el mapa de vecinos de dicha categoría. Figura 3.

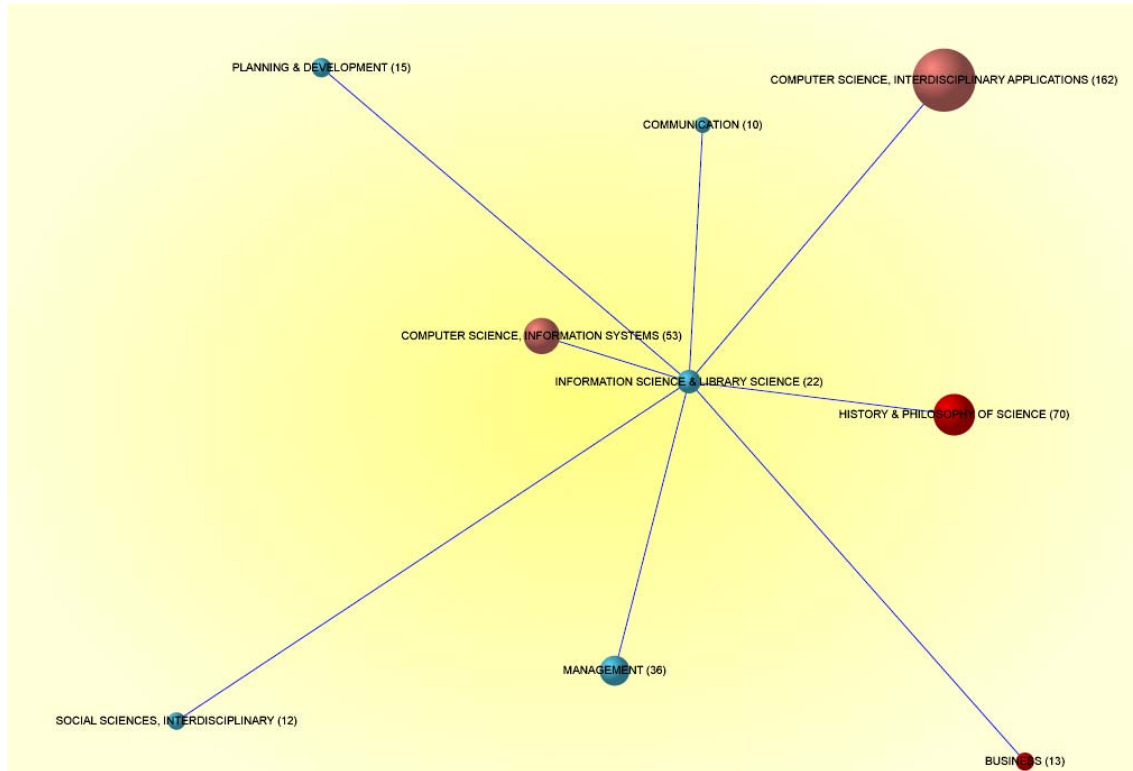


Figura 3 Mapa de Vecinos. Social Sciences

Al igual que ocurre en los mapas de cocitación de categorías o de segundo nivel, de los que proceden, cada nodo de la red representa a una categoría *ISI*. Como es típico en este tipo de redes, la representación ocupa al máximo el espacio disponible.

La principal característica de este tipo de mapas es que están formados por una red egocéntrica o en forma de estrella, donde el nodo objeto de estudio siempre se sitúa en el centro y el resto orbitan o circundan a éste con una mayor o menor distancia, dependiendo de la intensidad de sus relaciones. Por ejemplo, para el caso que nos ocupa, las categorías más cercanas y por tanto con las que está más estrechamente relacionadas son, por este orden: *Computer Sciences & Information Systems*, *Communication*, *History & Philosophy of Science*, *Management*, *Computer Sciences & Interdisciplinary Application*, *Planning & Development*, *Business* y *Social Sciences, Interdisciplinary*. Otra característica de estos mapas es que el nodo o categoría central agrupa en torno a sí, a aquellos nodos o categorías con los que mantiene una mayor relación, en cuanto a intercambio de fuentes se refiere, independientemente del área temática o clase *ANEP* a la que pertenezca. Por lo que resulta muy fácil detectar las agrupaciones entre categorías y su intensidad, independientemente de cualquier tipo de clasificación externa.

6. Conclusiones

Hemos demostrado que es posible construir mapas de grandes dominios, en este caso geográficos, por medio de la cocitación de clases o categorías. Por tanto, nuestra propuesta de uso de la cocitación de clases y de categorías, como entidades de cocitación y unidades de medida para la generación de visualizaciones esquemáticas y análisis de grandes dominios científicos, se constituye en una nueva herramienta para la generación de grandes mapas del conocimiento, y con ellos confeccionar el atlas de la ciencia

Los mapas desarrollados mediante esta nueva técnica, pueden ser muy útiles para la comparación de dominios similares con ubicaciones geográficas, temáticas o institucionales distintas. Al mismo tiempo que, dentro de un mismo dominio, pueden detectar e identificar la aparición y evolución de frentes de investigación mediante su representación en secuencias temporales.

Para el análisis de dominios, proponemos como orden de prelación o jerarquía los mapas de clases, seguidos de los de categorías y en último lugar los mapas de vecinos[60-61]. La continuación natural sería que a los últimos les siguieran mapas de cocitación de revistas y de cocitación de autores, por este orden. Estamos trabajando ya en ello, y en próximos trabajos, informaremos sobre la metodología y las técnicas para su construcción.

Es cierto que, por ahora, los mapas construidos mediante la metodología propuesta, no son más que representaciones esquemáticas off-line. Pero también lo es que hemos desarrollado una serie de herramientas, que nos permiten la generación automática de dichos mapas, para un dominio como España, en muy poco tiempo. No obstante, no hay que olvidar que sólo es necesaria la regeneración de los mapas cuando se incorpora nueva información a la base de datos, hasta ahora, con una periodicidad anual. A pesar de ello, no descartamos que en un futuro, puedan ser generados de forma dinámica.

Para todo aquel que lo desee, y para completar los mapas y comentarios aquí expuestos, hay una “demo” o prototipo temático de los dominios geográficos correspondientes a España y Chile, disponibles para toda persona que esté interesada en ellos en www.atlasofscience.net.

Bibliografía

1. Framework of a visualization system, [On-line], http://www.siggraph.org/education/materials/HyperVis/abs_con1/main.htm [Consultado en: 14/04/2003].
2. Thomson ISI, *ISI Web of Science* [Web page], Available at: <http://wos.cica.es/> [Consultado en: 2003].
3. E. Garfield, I.H. Sher, y R.J. Torpie, The use of citation data in writing the history of science, Philadelphia, Institute for Scientific Information, 1964.
4. L. B. Doyle, Semantic roadmaps for literature searchers, *Journal of the Association for Computing Machinery*, 8, (1961) No. 4, 553-578.
5. E. Garfield, Citation indexes in sociological and historical research, *American Documentation*, 14, (1963) No. 4, 289-291.
6. J. D. Price Derek, Networks of scientific papers, *Science*, 149, (1965) 510-515.
7. H. Small, Co-citation in the scientific literature: a new measure of the relationship between two documents, *Journal of the American Society for Information Science (JASIS)*, 24, (1973) 265-269.
8. V. Marshakova, System of document connection based on references, *Nauchno-Tekhnicheskaya Informatsiya*, Series II, (1973) No. 6, 3-8.
9. H. Small and B. C. Griffith, The structure of scientific literature, I: identifying and graphing specialities, *Science Studies*, 4, (1974) 17-40.
10. B. C. Griffith, H. Small, J. A. Stonehill, y S. Dey, The structure of scientific literature, II: toward a macro and microstructure for science, *Science Studies*, 4, (1974) 339-365.
11. S. Aaronson, The footness of science, *Mosaic*, 6, (1975) No. march-april, 22-27.
12. G. Salton and D. Bergmark, A citation study of computer science literature, *Professional Communication, IEEE Transaction*, PC-22, (1979) 146-158.
13. E. Garfield, Introducing the ISI Atlas of Science: Biochemistry and molecular biology, 1978-80, *Current Contents*, (1981) No. 42, 5-13.
14. E. Garfield, Introducing the ISI Atlas of Science: Biotechnology and molecular genetics, 1981/82 and bibliographic update for 1983/84, *Current Contents*, (1984) No. 41, 3-15.
15. E. Garfield, The encyclopedic ISI-Atlas of Science launches 3 new sections: bochemistry, inmunology, y animal and plant sicences, *Current contents*, (1988) No. 7, 3-8.
16. L. S. Seiden and D. R. Swanson, ISI Atlas of Science: Pharmacology 1987, Vol 1 - Inst-SCI- Informat, *Library Quarterly*, 59, (1989) No. 1, 72-73.
17. H. Small, The relationship of information.science to the social -sciences - a co-citation analysis, *Information Processing & Management*, 17, (1981) No. 1, 39.
18. H. Small and E. Garfield, The geography of science: disciplinary and national mappings, *Journal of Information Science*, 11, (1985) No. 4, 147-159.
19. H. Small and E. Sweeney, Clustering the science citation index using co-citations.1. A comparison of methods, *Scientometrics*, 7, (1985) No. 3-6, 391.
20. H. Small, E. Sweeney, y E. Greenlee, Clustering the science citation index using co.citations.2. Mapping science, *Scientometrics*, 8, (1985) No. 5-6, 321-340.
21. H. Small, Macrolevel changes in the structure of cocitation clusters: 1983-1989, *Scientometrics*, 26, (1993) No. 1, 5.
22. H. Small, A SCI-MAP case-study: building a map of aids research, *Scientometrics*, 30, (1994) No. 1, 229.
23. H. Small, Visualizing science by citation mapping, *Journal of the American Society for Information Science (JASIS)*, 50, (1999) No. 9, 799-813.
24. H. Small. Charting pathways through science: exploring Garfield's vision of a unified index to science. En B. Cronin and H. B. Atkins (Eds.), *The web of knowledge:*

- a festschrift in honor of Eugene Garfield (pp.). New Jersey: Information Today, 2000.
25. K. Börner, C. Chen, y K. W. Boyack, Visualizing knowledge domains, *Annual Review of Information Science & Technology*, 37, (2003) 179-255.
 26. R. R. Braam, H. F. Moed, y A. F. J. v. Raan, Mapping os Science by combined co-citation and word analysis. I: estructural aspects, *Journal of the American Society for Information Science (JASIS)*, 42, (1991) No. 4, 233-251.
 27. R. R. Braam, H. F. Moed, y A. F. J. v. Raan, Mapping os Science by combined co-citation and word analysis. II: dynamical aspects, *Journal of the American Society for Information Science (JASIS)* , 42, (1991) No. 4, 252-266.
 28. X. Lin, D. Soergel, y G. Marchionini. A self-organizing semantic map for information retrieval, En: Proceedings of the Fourteenth Annual International ACM/SIGIR Conference Research and Development in Information Retrieval, (262-269), Chicago, 1991.
 29. B. Hjørland and Albrechtsen, Toward a new horizon in information science: domain analysis, *Journal of the American Society for Information Science (JASIS)*, 46, (1995) No. 6, 400-425.
 30. H.D. White and K.W. McCain, Visualization of literatures, *Annual Review of Information Systems and Technology (ARIST)*, 32, (1997) 99-168.
 31. H.D. White and K.W. McCain, Visualizing a discipline: an author co-citation analysis of information science, 1972-1995, *Journal of the American Society for Information Science (JASIS)*, 49, (1998) No. 4, 327-355.
 32. E. Garfield, Mapping the world of science (at the 150 Anniversary Meeting of the AAAS, Philadelphia, PA), [On-line], <http://www.garfield.library.upenn.edu/papers/mapsciworld.html> [Consultado en: 07/05/2003], 1998.
 33. H. D. White, X. Lin, y K. W. McCain. Two modes of automated domain analysis: multidimensional scaling vs. Kohonen feature mapping of information science authors, En: Proceedings of the Fifth International ISKO Conference , (57-61), Würzburg, Ergon Verlag, 1998.
 34. Y. Ding, G.G. Chowdhury, S. Foo, y W. Qian, Bibliometric information retrieval systems (BIRS): a web search interface utilizing bibliometric research results, *Journal of the American Society for Information Science (JASIS)*, 51, (2000) No. 13, 1190-1204.
 35. R. K. Merton. On the Garfield imput to the sociology of science: a retrospective collage. En B. Cronin and H. B. Atkins (Eds.), *The web of knowledge: a festschrift in honor of Eugene Garfield* (pp.). New Jersey: Information Today, 2000.
 36. H. D. White. Toward ego-centered citation analysis. En B. Cronin and H. B. Atkins (Eds.), *The web of knowledge: a festschrift in honor of Eugene Garfield*, New Jersey: Information Today, 2000.
 37. E. C. M. Noyons, H. F. Moed, y M. Luwel, Combining mapping and citation analysis for evaluative bibliometric purposes: a biliometric study, *Journal of the American Society for Information Science (JASIS)*, 50, (1999) No. 2, 115-131.
 38. R. K. Buter and E. C. M. Noyons, Improving the functionality of interactive bibliometric science maps, *Scientometrics*, 51, (2001) No. 1, 55-67.
 39. E. C. M. Noyons, Bibliometric mapping of science in a science policy context, *Scientometrics*, 50, (2001) No. 1, 83-98.
 40. C. Chen and R. J. Paul, Visualizing a knowledge domain's intellectual structure, *Computer*, 34, (2001) No. 3, 65-71.
 41. C. Chen, R. J. Paul, y B. O'keefe, Fitting the jigsaw of citation: information visualization in domain analysis, *Journal of the American Society for Information Science and Technology (JASIST)*, 52, (2001) No. 4, 315-330.
 42. Y. Ding, G.G. Chowdhury, y S. Foo, Bibliometric cartography of information retrieval research by using co-word analysis, *Information Processing & Management*, 37, (2001) 801-817.
 43. P. Ingwersen and B. Larsen, Mapping national research profiles in social science

- disciplines, *Journal of Documentation*, 57, (2001) No. 6, 715-740.
44. V.P. Guerrero Bote, F.d. Moya Anegón, y V. Herrero Solana, Document organization using Kohonen's algorithm, *Information Processing & Management*, 38, (2002) 79-89.
 45. H. D. White, J. Buzydlowski, y X. Lin. Co-cited author maps as interfaces to digital libraries: designing Pathfinder Networks in the humanities, En: IEEE International Conference on information visualization, (25-30), 2000.
 46. X. Lin, H. D. White, y J. Buzydlowski, Real-time author co-citation mapping for online searching, *Information Processing & Management*, (In press) 18 p.
 47. J. Buzydlowski, H. D. White, y X. Lin, Term co-occurrence analysis as an interface for digital libraries, *Lecture Notes in Computer Science Series*, 2539, (2002) 133-144.
 48. H.D. White, Author-centered bibliometrics through CAMEOs: characterizations automatically made and edited online, *Scientometrics*, 51, (2001) No. 3, 607-637.
 49. H. Small, Paradigms, citations and maps of science: a personal history, *Journal of the American Society for Information Science and Technology (JASIST)*, 54, (2003) No. 5, 394-399.
 50. H. D. White, Pathfinder networks and author cocitation analysis: a remapping of paradigmatic information scientist, *Journal of the American Society for Information Science and Technology (JASIST)*, 54, (2003) No. 5, 423-434.
 51. C. Chen and J. Kuljis, The rising landscape: a visual exploration of superstring revolutions in physics, *Journal of the American Society for Information Science and Technology (JASIST)*, 54, (2003) No. 5, 435-446.
 52. S.A. Morris, G. Yen, Z. Wu, y B. Asnake, Time line visualization of research fronts, *Journal of the American Society for Information Science and Technology (JASIST)*, 54, (2003) No. 5, 413-422.
 53. K.W. Boyack and K. Börner, Indicator-assisted evaluation and funding of research: visualizing the influence of grants on the number and citation counts of research papers, *Journal of the American Society for Information Science and Technology (JASIST)*, 54, (2003) No. 5, 447-461.
 54. E.R. Tufte, *Envisioning information*, Cheshire, Graphics Press, 1994.
 55. E.R. Tufte, *The visual display of quantitative information*, Cheshire, Graphics Press, 2001.
 56. Thomson ISI, *ISI Journal Citation Reports* [Web page], Available at: <http://isiknowledge.com> [Consultado en: 2003].
 57. Vladimir Batagelj and Andrej Mrvar, *Pajek*, Networks/Pajek: Package for large network analysis, 2003.
 58. T. Kamada and S. Kawai, An algorithm for drawing general undirected graphs, *Information Processing Letters*, 31, (1989) No. 1, 7-15.
 59. W3C, *Extensible Markup Language (XML)* [Web page], Available at: <http://www.w3.org/XML/> [Consultado en: 2003], 2003).
 60. F. de Moya-Anegón, B. Vargas-Quesada, V. Herrero-Solana, Z. Chinchilla-Rodríguez, E. Corera-Álvarez y F. J. Muñoz-Fernández. A new technique for building maps of large scientific domains based on the cocitation of classes and categories. *Scientometrics*, 61, (2004) No. 1, 129-145.
 61. F. de Moya-Anegón, B. Vargas-Quesada, Z. Chinchilla-Rodríguez, V. Herrero-Solana, E. Corera-Álvarez y F. J. Muñoz-Fernández. Domain analysis and information retrieval through the construction of heliocentric maps based on ISI-JCR category cocitation. *Information Processing & Management*, 41, (2005) No 6, 1520-1533.