

ANÁLISIS DE REDES SOCIALES A PARTIR DE RECURSOS WEB Y DE BASES DE DATOS ESPECIALIZADAS EN LITERATURA CIENTÍFICA

*Adilson Luiz Pinto**

Departamento de Biblioteconomía de la UFMT. Rondonópolis, Mato Grosso.

*José Antonio Moreira González***

Departamento de Documentación. Universidad Carlos III de Madrid.

*Alexandre Oliveira de Meira Gusmão****

Departamento de Biblioteconomía de la UFMT. Rondonópolis, Mato Grosso.

Resumen: Se describen los resultados de una reflexión basada en los fundamentos y tendencias científicas de las Redes Sociales, cuya estructura se aborda desde los principios y actuaciones de los colegios invisibles y desde el capital social; mientras que, por otra parte, se exploran las técnicas de generación de mapas de relaciones utilizando para ello programas de software libre (Pajek, CiteSpace y TouchGraph). En la práctica seguida, se han empleado dichos programas para el estudio de casos, observándose que cada plataforma tuvo un comportamiento diferente, causado en las características propias de cada recurso informático.

Palabras clave: Red social; software para análisis de relaciones; colegios invisibles; capital social.

Title: SOCIAL NETWORK ANALYSES FROM RESOURCES WEB AND DATABASES SPECIALIZED IN SCIENTIFIC LITERATURE.

Abstract: This work tried to describe the schools and scientific bases of the Social Network, with their structures made under the invisible colleges and social capital; focusing also in the explanation of the techniques in creation of maps for relationship through free softwares (Pajek, CiteSpace and TouchGraph). The practicum development has been based to turn the mentioned programs into application of cases, where each platform has shown a different behavior, caused by the self characteristics of each and every informatic resource.

Keywords: Social network; software for analysis of relations; invisible college; social capital.

1 EL CONCEPTO DE RED SOCIAL. SU EVOLUCIÓN

La idea de Red Social comprende aquel aspecto de la relación y cooperación entre los integrantes de un conjunto por el que se conectan horizontalmente, bien sea de manera directa o a través de los miembros que les son próximos.

* adilson.luz2208@hotmail.com

** jamore@bib.uc3m.es

*** aomgusmao@hotmail.com

Este hecho se manifiesta en la forma de una malla constituida por múltiples secuencias, que puede esparcirse indefinidamente hacia todos los lados, sin que ninguno de sus nodos pueda ser considerado principal o central, ni representante de los demás. No hay un “comandante”, lo que hay es una voluntad colectiva de realizar y alcanzar determinados objetivos en conjunto.

Durante las últimas décadas, el concepto de red social y los análisis de estas relaciones han sido desarrollados como una de las vías más prometedoras para medir la estructura social y las cooperaciones existentes en ella. Su fundamentación, sin embargo, se debió a la introducción hecha por Jacob Moreno (1934), buscando consolidar la Sociometría, de la teoría matemática de grafos, que conoció una adaptación posterior con la aplicación ideada por Cartwright y Harary (1956), que se conoce como Teoría del Equilibrio Estructural.

En coincidencia temporal, Heider (1946) aportaba un estudio sobre el análisis de equilibrio en las relaciones, partiendo de una experiencia cognitiva. Determinó dos tipos de posibles cooperaciones: las originadas en la simpatía, o positivas, cuando las relaciones son obvias dentro de un universo; y aquellas derivadas de la desaprobación o negación que, en general, están formadas por casualidades que suceden dentro de una cooperación que no posee centralidad.

A partir de estos marcos iniciales, Scott (1991) fijó las bases de la Teoría de Redes Sociales, que se desarrollarían en la Universidad de Manchester en la década de 1950. Al frente de este proyecto se encontraban Max Gluckman (1954) y John Barnes (1954), quizás los exponentes más importantes en dicha Teoría.

Puede considerarse que su proceso de establecimiento y consolidación se logró a partir de tres bloques, el primero integra las teorías utilizadas y desarrolladas (Teoría de la Gestalt, Teoría Sociométrica, Dinámica de Grupos y Teoría de Grafos); el segundo, las nuevas escuelas científicas (Antropología estructural-funcional, Estructura de Harvard y Análisis de las Redes Sociales); y, finalmente, el de los creadores que participaron más activamente (Warner y Mayo, Homans, Gluckman, Mitchell, Barnes y Bott).

En su vertiente antropológica se hicieron estudios aplicados a la documentación médica y sus implicaciones (análisis prospectivo -longitudinal o transversal- de documentos clínicos y control de las historias clínicas con análisis de sus datos), así como a la normalización de las técnicas de tratamiento en grupo, e incluso, desde presuposiciones cognitivas, se dirigieron al control del tratamiento en escala de redes de pacientes (Litwin, 1997).

Otro sector de uso habitual ha sido el del mundo de la empresa, principalmente en lo relativo al control de grupos y de los procesos de producción, así como a la hora de identificar clusters empresariales (Grannovetter, 1985).

Estas perspectivas (antropológicas) se derivan directamente de los estudios iniciales sobre colegios invisibles y sobre el capital social, que siempre se tomaron desde el prisma de relación entre personas o instituciones.

Desde esta premisa, resulta esencial observar que el fenómeno de las Redes Sociales no presenta solamente un carácter evolutivo. Se trata de una disciplina que genera nuevo conocimiento desde la perspectiva de solución de situaciones opuestas (incluso de los momentos difíciles de superar) en el área de las Ciencias Sociales, por trabajar con acciones estructuradas y con los pares micro y macro y, al mismo tiempo, por explorar sobre todo la generación “cuantitativa” de datos, proceso en el que se fundamentan la Sociometría y la Teoría de grafos.

Debe de considerarse como contribución relevante el trabajo de Molina, Muñoz y Domenech (2002), en el que describieron la posibilidad de generar una Red Social para estudios de carácter métrico, utilizando la estructuración de las coautorías para identificar las relaciones de cooperación científica a nivel individual, de grupo y en red de autores.

Para fortalecer este aspecto cuantitativo de las Redes resulta esencial que se adopten los estudios métricos a la hora de fundamentar otros aspectos que muestran las relaciones, como las densidades, las intermediaciones y las cercanías, que son cálculos que determinan el grado de la centralidad existente. En definitiva, aplicar las Matemáticas a los estudios Sociales y Humanos.

Esta perspectiva se robustece si consideramos que el conocimiento de las Redes Sociales tiene un fundamento estadístico, puesto que resultan de la unión cooperativa de dos o más puntos, o simplemente por constituir una representación de la frecuencia diseñada en un mapa por inferencias similares (de autores, temáticas, revistas, documentos o instituciones científicas, o de acciones municipales o del control del crimen, entre otras actividades).

Las Redes Sociales determinan diversos comportamientos, en lo que respecta a su faceta antropológica, y cuentan con diversos movimientos desde los que se fundamentó su modelo actual. Dentro de las aportaciones teóricas que han confluído en los conceptos vigentes sobre Redes Sociales, se pueden destacar las provenientes de los colegios invisibles y del capital social, que se comentarán a continuación, así como su trayectoria y su fundamentación respecto a los estudios de relaciones.

1.1 El origen de las redes sociales: su base en los colegios invisibles

En 1660, a los pocos meses de la restauración del reinado de Carlos II de Inglaterra, un grupo de doce investigadores, entre los que se encontraban Robert Boyle y Christopher Wren, se reunieron en Londres con el objetivo de crear una sociedad dedicada a estudiar los mecanismos de la Naturaleza. Se enfrentaban a una ardua tarea, pues el mundo occidental vivía aún impregnado por la fe, la superstición y la magia, y estaba lejos de emplear la razón para explicar los fenómenos naturales y sociales. De aquella iniciativa surgió la *Royal Society* y, con ella, las primeras iniciativas para divulgar las ideas científicas (Lomas, 2006).

En actualidad más contemporánea Crane (1972) estudió la actuación de los colegios invisibles, denominación dada a las comunidades informales de científicos que trabajan en la misma temática y que intercambian informaciones. Es característica de estos colegios la interacción y la comunicación dentro de una red, que funciona como un forum educativo. En este sentido, beneficia tanto a los científicos expertos como a los recién llegados al trabajo en un determinado tema, que necesitan de informaciones aplicables, así como conocer algunas técnicas o donde se encuentra su fundamentación teórica.

Para el ámbito de los colegios invisibles se cuenta con la Sociometría, representación gráfica de base matemática y características cartesianas, que se situó como ciencia desde el momento en que fue capaz de estudiar una estructura social simultáneamente en su conjunto y en sus partes.

Se trata, por lo tanto, de una ciencia social que se ocupa de los individuos que conforman un grupo determinado, pero también de las relaciones que tienen con sus pares. Esta fundamentación estructural puede representarse mejor por los modelos que la simbolizan:

- *Estrella*: compuesta de diversas líneas que apuntan hacia un punto central que representa a un individuo envuelto por la relación con cada uno del resto de participantes, y donde las flechas representan relaciones tales como reciprocidad, ignorancia o rechazo.
- *Corriente*: se caracteriza por una serie de relaciones en fila, con un centro que intermedia pares.
- *Pares*: compuesto por una relación recíproca entre dos individuos.
- *Poder*: representado como una atracción entre grupos que se funden por consecuencia de uno de los individuos.
- *Aislamiento*: determinado por un asunto no escogido por muchos individuos, cuya interacción social no está consolidada, pero que es representable como parte del proceso de análisis Sociométrico.

Estos modelos son esencialmente aplicables a los análisis temporales, además tienen un funcionamiento orientado a establecer combinaciones dentro de la red, con finalidad de justificar líneas de investigación, conocer los planes pilotos de una institución, o el control de plagas, o el proceso seguido por un contagio o una influencia.

Por otra parte, el estudio de redes sociales se basa en otro marco teórico, denominado capital social, cuya función se expone a continuación.

1.2 El capital social como aspecto social de las redes

El concepto lo comenzó a aplicar Fischer (1959) para describir una calificación en administración de negocios, encaminando de esta manera los principios antropológicos hacia el interés empresarial.

Posteriormente, dos estudios tuvieron una excelente aceptación por parte de la crítica:

- I. Psicología cultural y del riesgo de aplicación social del capital para formar nuevas líneas de investigación, basada en la teoría de los colegios invisibles y de sus relaciones (Ponton, 1975).
- II. Economía como trampolín para la consolidación social de las redes, desde una orientación hacia la Sociología y sus vertientes económicas y políticas, además de la parte aritmética de las relaciones (Rose, 1961).

Coincidentemente, Pierre Bourdieu (1977) establecía la distinción entre el capital económico, su simbología, su cultura y su acción social, que causó gran impacto en la aplicación de la Red Social al ámbito de las oportunidades laborales, al desarrollo de las profesiones y a la gestión de contactos. Estas iniciativas se representaron siempre mediante herramientas de gestión.

Coleman (1988) sentó que existe una relación de complementación directa entre el capital económico (infraestructura, financiación), el capital humano (educación) y el capital social (relaciones de confianza). Este hecho hace que se potencie la unión entre el capital económico y el capital humano, en la medida en que aumenten las relaciones de confianza y de cooperación en la comunidad. En este trabajo, Coleman discute dos corrientes que coinciden de manera diferente con la acción social: la primera, considera al actor como sobre-socializado, gobernado por normas y sobre-determinado por el sistema; y la segunda, observa al actor como subsocializado, es decir, con pocos condicionamientos en la comunidad.

En 1995, Putnam, seguidor de Coleman, describe que el capital social se refiere a las características de la organización social, tales como la confianza, las normas y las redes, y que pueden mejorar la eficiencia de la sociedad si se favorecen las acciones coordinadas. En ese trabajo Putnam determinó categorías como:

1. *Confianza*: predicción de la conducta de un actor independiente. Su hipótesis parte de que la confianza es un concepto sin relevancia sociológica, circunscrita al orden emocional del individuo. Conviene recordar que tiene enormes implicaciones para la forma en que se generan socialmente las relaciones y en cómo se producen los intercambios.
2. *Normas*: implicaciones que señalaba Coleman sobre las externalidades, es decir, la posibilidad de alcanzar consecuencias negativas o positivas para un mismo actor y para los demás. La reciprocidad es la norma más relevante en términos de capital social. Se la considera de dos tipos: (i) reciprocidad equilibrada o específica; (ii) y generalizada o difusa.
3. *Redes*: en comunidades donde se puede esperar que la confianza no sea aprovechada, si no correspondida, es más probable que el intercambio continúe. En consecuencia, la reciprocidad generalizada es un componente de la generación de capital social ya que estimula la posibilidad de que los individuos cooperen entre sí para beneficio común.

2 INICIATIVAS DE GENERACIÓN GRÁFICA DE REDES SOCIALES

El análisis de una red se basa en la noción intuitiva de que estos patrones son características importantes en las vidas de los individuos que los exhiben. Los analistas de redes creen que el modo de vida de un individuo en gran parte depende de cómo esté enlazado en el tejido mayor de las conexiones sociales. Por otra parte, es creencia muy extendida que el éxito o la ausencia de asociaciones y de organizaciones dependen, a menudo, de modelar su estructura interna.

Partiendo de una visión de dependencias y de relaciones, se han desarrollado programas y *software* específicos que interpretan estas interacciones mediante teorías tomadas de las Matemáticas y de la Estadística, luego aplicadas a la Informática siempre desde una categorización que trabaja las informaciones en fracciones cartesianas.

Una de las funciones que los programas trabajan por detrás de sus interfaces son los lazos fuertes y débiles, para cuya identificación los programas utilizan, en general, algo parecido a una representación de poder a través de la frecuencia y de la intensidad con que se determinan las relaciones.

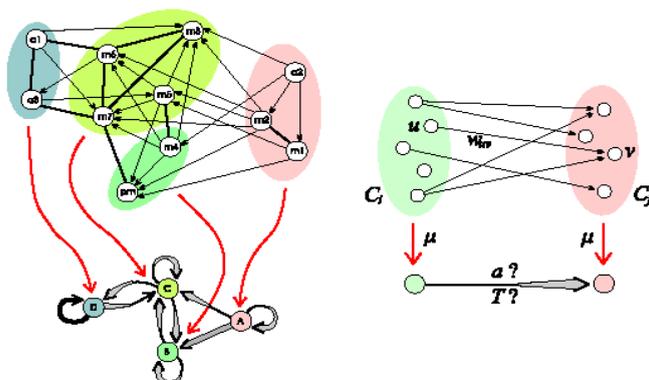


Figura 1. Relaciones posibles a través de programas de Red Social. Fuente: Batagelj y Mrvar (2003, p. 3).

Se puede identificar una red completa. Se deduce que en el primer bloque existe una gran intensidad de relaciones, con grupos caracterizados por sus interacciones internas, pero que al mismo tiempo también se relacionan externamente, demostrando algunos enlaces débiles (los que aparecen en los extremos), y otros de centralidad y poder (como es el caso del grupo en verde).

Otra función muy convencional es la identificar la estructura de relaciones habidas en el tiempo, determinadas en inferencias por año, décadas y hasta siglos.

Para la representación gráfica de las Redes utilizamos algún programa de *software* libre, con mucho prestigio en el mundo científico, como *Pajek*, *CiteSpace* y *Touch Graphic* buscando una mejor explicación de las funciones y buena aplicación práctica. Es importante destacar que estos programas sirven para analizar todo tipo de estructuras, aunque nuestro análisis tendrá la preocupación principal de explicar cómo funcionan estos recursos, cómo los aplicaremos en el universo científico, fuera de su uso social original.

2.1 Pajek <<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>>

Pajek es un programa que sirve para analizar y visualizar grandes redes, a partir de colaboraciones simbolizables en una malla similar a una tela de araña, con relaciones horizontales. El diseño de Pajek se basa en experiencias alcanzadas al desarrollar estructuras de datos gráficos. Su función es generar redes a través de algoritmos cartesianos y del X-gráfico, ejecutados a partir de las relaciones y fundamentación de STRAN, RelCalc, Draw, Energ y SGML desde la descripción gráfica del lenguaje NetML (Batagelj y Mrvar, 2003).

Este programa tiene como fin principal desarrollar grandes redes sociales a partir de redes pequeñas, asimismo, visualizar gráficamente herramientas y técnicas científicas basadas en la Teoría de Grafos, e identificar los cálculos de densidad, intermediación y proximidad científica de las cooperaciones.

Su representación viene moldeada desde la Sociometría y está trabajada desde una perspectiva de visualización en dos ambientes (global y local), con estructuras jerárquicas, de fracciones cortadas, de contextos y de reducción.

La estructura reproductora de las redes aborda algunos campos esenciales para la configuración gráfica y estadística de la Red Social. Dentro de estos campos destacamos:

- **Networks:** que tiene como objeto principal determinar los vértices y las líneas. Su extensión de lectura se hace desde **.net**. Con este campo la red se puede presentar en fichero de entrada en *arcs/edges* (ejemplo: 1, 2 - línea a partir de la 1 a 2) o a partir de *arcslists/edgeslists* (ejemplo 1, 2 y 3 - línea a partir de 1 a 2 y a partir de 1 a 3) o en formato de matriz a través de otros programas como UNICET o GEDCOM. La información adicional para el dibujo de la red también se puede incluir en fichero de entrada. Esto se explica en las exportaciones de la sección a EPS/SVG/VRML.
- **Partitions:** tiene como característica determinar la relación que posee cada vértice. Su extensión por defecto es en **.clu**.
- **Permutations:** funciona como un reordenador de vértices. Su extensión es **.per**.
- **Clusters:** es el subconjunto de los vértices, como una clase de la partición. Su extensión está clasificada por jerarquías de **.cls**.
- **Hierarchies:** son los vértices jerárquicos. Su función parte del *Root* (la raíz) y tiene por característica dos subgrupos - g1 y g2. g2 es una hoja - clusters con vértices v5, v6 y v7. g1 tiene dos subgrupos - g11 y g12. La extensión por defecto se hace en **.hie**.

Ejemplo: Root

```
g1 g2
g11 g12 v5, v6, v7
v1, v2 v3, v4
```

- **Vectors:** determina una característica numérica cierta (número verdadero) para cada vértice. Su extensión es en **.vec**.

La relación que se genera para una configuración visual es cartesiana y su estructura tiene que estar representada por una descripción de los participantes y su cantidad de operaciones, como el ejemplo a seguir:

*Vértices 11

```
1 " CSIC "
2 " Univ Carlos III "
3 " Univ Coimbra - Portugal "
4 " Univ La Coruña "
5 " Univ Navarra "
6 " Univ Oviedo "
7 " Univ Politécnica Madrid "
8 " Univ Sevilla "
9 " Univ Valladolid "
10 " Univ Zaragoza "
11 " Centro Nac Inv Metalúrgicas "
```

*Matrix

0	2	0	0	1	0	1	0	0	0	0
2	0	2	2	2	3	6	1	1	1	2
0	2	0	0	0	0	1	0	0	0	1
0	2	0	0	0	1	1	0	0	0	0

1	2	0	0	0	0	1	0	0	0	0
0	3	0	1	0	0	1	1	1	1	0
1	6	1	1	1	1	0	0	0	0	1
0	1	0	0	0	1	0	0	1	1	0
0	1	0	0	0	1	0	1	0	1	0
0	1	0	0	0	1	0	1	1	0	0
0	2	1	0	0	0	1	0	0	0	0

Para generar los gráficos es importante tener en cuenta que puede desarrollarse a partir de *arcs/edges* y *arcslists/edgeslists*, en un proceso que puede generarse automáticamente, utilizando otros programas para ejecutar la transferencia de extensiones desde el formato *txt* hacia una versión *.net*, que es la ideal para la representación gráfica. Uno de los programas que realiza esta función es el *CiteSpace*, software que se comentará a continuación.

2.2 CiteSpace <<http://cluster.cis.drexel.edu/~cchen/citespace/>>

CiteSpace es un programa desarrollado por Chaomei Chen (2006), de la Universidad de Drexel, para visualizar la evolución de una red a través de un número de intervalos temporales, con representación de sus nudos y frecuencias. *CiteSpace* es un programa diseñado en *Java* para analizar exclusivamente las cocitaciones, para lo que utiliza técnicas subyacentes al número de repeticiones como elemento para determinar las mayores incidencias en citas.

Actualmente concentra sus fuerzas en extraer informaciones de los productos del *Institute for Scientific Information* y del *PubMed*, aunque la inventora está preocupada en expandir su aplicación técnica automática. Una manera común de utilizar esta herramienta es para los intervalos de tiempo en segmentos menores y estudiar y así estudiar cómo se comporta individualmente cada nudo en la red en factores temporales.

La función de este programa es identificar las relaciones entre los autores que más producen, los términos más utilizados por los autores, los autores más citados, los documentos más citados y las revistas más citadas. Parece evidente que con este recurso se pueda identificar la popularidad de los autores y también el consumo de información hecho por los autores.

Para comprender el funcionamiento estadístico y gráfico de *CiteSpace* es fundamental entender los conceptos que generan sus relaciones. También es importante tener en cuenta que este programa sigue una rígida estructura, que establece campos automáticos y su labor está en analizar redes de producciones científicas, sin embargo desde que se genera un documento fuente con las mismas características puede ser visualizado para otros intereses, como análisis de ontológico y de inteligencia artificial.

A continuación identificaremos los conceptos claves de este programa:

- **Betweenness centrality:** métrica del nodo de la red que estipula la medida de la trayectoria arbitraria más corta de la red, conocida también como intermediación.
- **Burst terms:** contenido de las frases o de palabras extraídas del título, del resumen o de otros campos de un documento bibliográfico, con la frecuencia de apariciones del término.
- **Citation:** referencias utilizadas para la representación gráfica de la red. Son las fuentes primarias del consumo de información.

- **Citation half-life:** número de citaciones que recibe una publicación desde su aparición por los años de vida.
- **Citation tree-rings:** círculos que determinan la representación de una serie temporal de citaciones, que siempre están caracterizadas por un análisis hecho de fuera hacia dentro. Siempre el meollo es proporcional a las citaciones en el año correspondiente.
- **Cluster view:** una red se visualiza mediante un algoritmo modificado por la colocación del centroide.
- **Co-authors:** representación de los nombres de los autores, independientemente de la cantidad de ellos que firmen el trabajo. Este campo es fundamental para identificar a los autores más productivos.
- **Co-citation:** representación en la que se identifican los consumos de información por parte de los autores. Figuran en ella los autores y documentos consultados, y las revistas.
- **Color map:** espectro de los colores usados por CiteSpace para representar el orden temporal de observaciones.
- **Expectation Maximization clustering:** son los centroides que surgen por poseer cualidades tales como citaciones, período en que son hechas, y la intermediación. El uso de cualidades temporales puede ayudar a visualizar las tendencias que vienen emergiendo.
- **Pathfinder network scaling:** algoritmo del escalamiento de la red que evita los acoplamientos que violan la desigualdad del triángulo. Condiciona la simplificación de una red al conservar los acoplamientos salientes y sus trayectorias.
- **Spotlight:** son las redes visualizadas que se sitúan fuera de aquellos acoplamientos que no conectan puntos giratorios.
- **Thresholds:** son criterios de selección, donde los artículos deben tener medidas sobre valores de tamaño para poder ser procesada y visualizada.
- **Time-zone view:** visión restringida en la que el movimiento de nodos se limita a las zonas de tiempo verticales que se corresponden con la época de su publicación.
- **Turning points:** nodos de alta intermediación (> 1.00). Los nodos tienden a ser críticos en transiciones de la red desde una franja de tiempo a otra.

La estructura que CiteSpace utiliza para sus análisis viene modulada por los campos de la *Web of Science* y de *PubMed*, cuyos contenidos se definen en cinco bloques. En el primer bloque se representa a los autores que firman los trabajos. Este aspecto es analizado en el tópico *Citing Authors*. El segundo bloque de análisis está contenido en las representaciones de los *Citing Terms*, que extrae informaciones provenientes del título, de los descriptores bibliográficos y del resumen. Los *Cited Authors*, *Cited Documents* y *Cited Journals* son generados por el bloque tres, donde se identifica todo el consumo de información utilizado por los autores que firman los trabajos. Es importante mencionar que el programa identifica para cada análisis los fragmentos y sus funciones, para no tener una desmedida recopilación de datos ni solapamientos. Finalmente, el cuarto bloque genera una serie temporal con los espacios de los nodos, identificando los colores de los mapas, las frecuencias de los grupos y subgrupos. Para la temporalidad se utiliza el bloque cinco, donde se efectúa un análisis de frecuencia por año.

2.3 TouchGraph <<http://www.touchgraph.com/>>

TouchGraph es un programa reciente desarrollado por la Compañía SourceForge.net, y que está disponible en libre acceso en Internet.

Su objetivo es determinar relaciones sociales dentro de la red mundial. Extrae informaciones de las páginas *web* e identifica cuales son sus hiperenlaces. Su estructura está diseñada para trabajar con Google, con Amazon y con páginas de revistas de acceso libre en Internet.

TouchGraph ofrece la misma solución gráfica que algunos motores de búsqueda como, por ejemplo, Teoma, Webbrain y Kartoo. Pero la estructura del TouchGraph está muy relacionada con el motor Vivisimo, capaz de determinar internamente una codificación entre resultados similares, de correlaciones y de sinónimos, y de las informaciones exactas buscadas.

Como herramienta de diagramas, el programa muestra relaciones neurológicas entre fragmentos de textos. Los usuarios pueden extraer porciones relevantes de literatura publicada y almacenarlas junto con los enlaces que indican su relación. El programa utiliza una función NeuroScholar que proporciona un ambiente conveniente para que el usuario pueda anotar los datos y para navegar usando el interfaz de red de TouchGraph.

A través de los hiperenlaces, TouchGraph puede analizar de una sola vez una media de 110 URLs, a partir de sus descriptores, sus títulos y las categorías DMOZ.

Cumple también la función de encontrar "camino de semejanza", es decir, sitios *web* parecidos pero que no aparecen como conectados. Puesto que todos los nodos que son recuperados en las distintas peticiones distan como máximo dos "vínculos de semejanza", si se añaden dos URLs en este mini formulario se puede saber si existe un camino de cuatro vínculos entre esos dos.

Finalmente, es importante mencionar que los gráficos no pueden alterarse, debido a ser una forma de representar gráficos basada en la Sociometría, por lo que mide siempre un valor radial, desde el que puede determinarse cuánto nodos se deben identificar en torno a un nodo concreto.

3 REPRESENTACIÓN PRÁCTICA DE LAS REDES SOCIALES

Las Redes Sociales tienen carácter teórico, en especial si consideramos su vocación original y su aplicación al campo de la Antropología. Sin embargo, si seguimos su variante exploratoria en Medicina podremos distanciarnos un poco de la visión antropológico-social, más aún si tenemos en cuenta algunas técnicas adquiridas de la Informática, que nos permiten observar cómo las investigaciones de esta última tienen una vocación de capital social que el resto de campos no contemplan, principalmente porque no generan gráficos desde los que se puedan visualizar las constantes relaciones.

En esta parte utilizaremos tres estudios de caso, enfocados, el primero, desde la Informática, desarrollando técnicas estadísticas, en el segundo, y, por fin, dando en el tercero una visión correspondiente a las Ciencias de la Información y Documentación.

3.1 Estudio de caso

El primer estudio de caso se aplicó a las relaciones existentes entre los investigadores que publican sus trabajos científicos en las plataformas *Francis* y *Social Science Citation Index*. El objeto de estudio fueron los docentes de universidades brasileñas y españolas

1. **Brasil:** con dos relaciones intensas, (1) entre la Universidad Federal de Minas Gerais y las Universidades de Brasilia, Federal Fluminense y de São Paulo; (2) y la relación de la Universidad de São Paulo con la Universidad de Brasilia.
2. **España:** con cuatro relaciones notables, (1) entre la Universidad de Granada y la Universidad de Extremadura; (2) entre la Universidad de Alcalá de Henares y la Universidad de Murcia; (3) entre la Universidad Complutense de Madrid y las Universidades de Zaragoza, Alcalá de Henares y Carlos III de Madrid; y (4) la relación entre la Universidad Carlos III de Madrid con las Universidades de Granada, Alcalá de Henares, Zaragoza, Murcia y Complutense de Madrid.

Otro aspecto relevante viene determinado por la cooperación científica entre las Universidades y los Centros de Investigación de ambos países. Esta cooperación fue constante entre la Universidad de São Paulo y la Universidad Federal de Minas Gerais con el IBICT en el caso brasileño; y la relación de la Universidad Complutense de Madrid con el CSIC/CINDOC para el caso español.

Las instituciones que presentan una centralidad más alta en cooperación científica fueron la Universidad Federal de Minas Gerais y la Universidad Carlos III de Madrid, principalmente porque se relacionaron con Fundaciones, Industrias y con el resto de las universidades brasileñas y españolas, junto a la presencia de centros de otros países como la Universidad de la Habana, la Universidad-Aix Marseille III, la Clark Atlanta University, la University Texas Austin y la Universidad de Stuttgart.

Dentro del análisis ha sido posible calcular grafos como el de densidad, la intermediación y la cercanía.

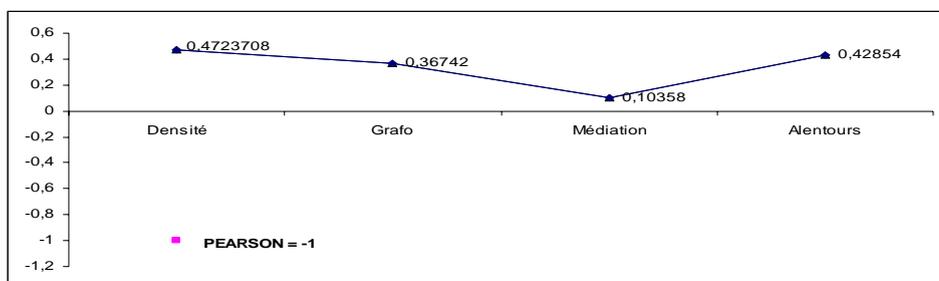


Figura 3. Resultado de centralidad de la muestra.

Estos cálculos son referencias a la hora de identificar el grado de centralidad existente en el análisis. En la muestra puede observarse también el coeficiente de correlación de Pearson. En la suma total se parecía una baja centralidad, con índices inferiores a 1 punto porcentual, por lo que dicho coeficiente se mostró con un valor de -1.

Este valor se deriva de la existencia de relaciones intensas, ya mencionadas anteriormente, pero que no evitan el hecho de que, en general, haya un escaso número de acciones de cooperación, y de que, cuando existen, sean de baja intensidad, muchas veces con un valor agregado de interacción inferior a los 0,5 puntos, ya que los vectores que alteran la cantidad inicial de relaciones, de acuerdo con el formato utilizado por el programa, permiten el cálculo fraccionado.

3.3 Análisis de los documentos citados en la Web of Science

Como segundo estudio de caso, averiguamos los documentos más citados dentro de las temáticas de Bibliometría y Cienciometría en las plataformas de la *Web of Science*, relativos al período abarcado entre los años 1975 y 2006.

Este caso práctico está enfocado al consumo de información relativo a las citas utilizadas por los autores. El análisis de documentos citados se hizo utilizando *Citespace*, obteniéndose diagnósticos de centralidad (mapa pequeño) y de frecuencia (mapa de la derecha).

Para entender mejor los mapas, se recomienda visualizar los nodos de tonalidad rosa, que significan el poder de centralidad; los nodos de tonalidad azul que representan la frecuencia; y los nodos en los que se mezclan tonalidades (azules internamente, con una línea rosa externa) que los destacan como documentos de fuerte centralidad y frecuencia.

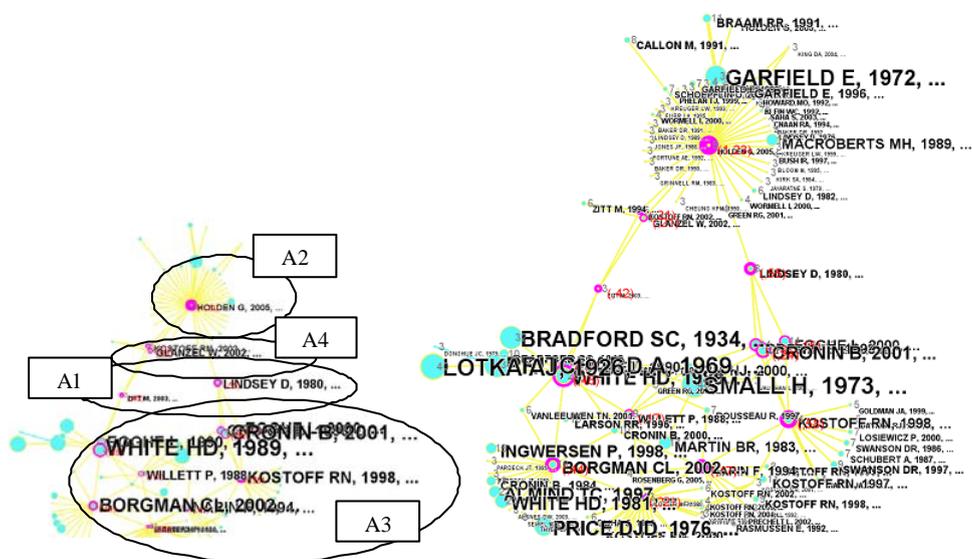


Figure 4. Proyección de estudios en CiteSpace: análisis del consumo de información en Web of Science.

Respecto a la centralidad debe destacarse el papel de Lindsey y Zitt (A1), cuyos documentos sirven de conexión entre dos grandes grupos. Estos documentos también obtienen una cantidad de frecuencia respetable. El segundo análisis, dentro de este mapa, habla de la centralidad total Holden (A2) en uno de los grupos, donde se encuentran citados otros 20 documentos.

El análisis referente al documento de Holden es considerado más importante por dos motivos: (i) se trata de un documento no demasiado citado (9 frecuencias), pero que sin embargo está relacionado con los documentos teóricos de relevancia; (ii) se trata además de un documento muy reciente, tal vez el más importante de 2005, cuando se publicó, que contiene la fundamentación de mayor relieve para las teorías que basan los estudios métricos.

Se observa otro caso de centralidad en los documentos que reciben más citas, como el de White, Cronin y Kostoff (1998), y el de Willett y Borgman (A3), que forman parte de un universo de cooperación científica muy representativa, como se destaca en la parte inferior del mapa menor.

Un último resultado de centralidad (A4) se deriva de los documentos intermediarios, cuya importancia es vital para la existencia gráfica del grupo A2, identificado en la parte superior del mapa.

A diferencia del análisis de centralidad, alejado de averiguar cuáles son los documentos más citados, la aplicación de parámetros de frecuencia si persigue ese objetivo. Alcanzándose los 35 documentos más utilizados en la temática de estudios métricos.

Ranking	Documentos Citados	Cantidad
1	LOTKA AJ, 1926, J WASHINGTON ACADEMY, V16, P317	58
2	PRICE DJD, 1963, LITTLE SCI BIG SCI, P62	48
3	PRITCHARD A, 1969, J DOC, V25, P348	46
4	BRADFORD SC, 1934, ENGINEERING-LONDON, V137, P85	43
5	GARFIELD E, 1979, CITATION INDEXING	42
6	SMALL H, 1973, J AM SOC INFORM SCI, V24, P265	41
7	WHITE HD, 1989, ANNU REV INFORM SCI, V24, P119	35
8	GARFIELD E, 1972, SCIENCE, V178, P471	32
9	PRICE DJD, 1965, SCIENCE, V149, P510	31
10	PRICE DJD, 1976, J AM SOC INFORM SCI, V27, P292	31
11	EGGHE L, 1990, INTRO INFORMETRICS Q	29
12	WHITE HD, 1981, J AM SOC INFORM SCI, V32, P163	29
13	CRONIN B, 1984, CITATION PROCESS ROL	28
14	NARIN F, 1976, EVALUATIVE BIBLIOMET	27
15	PRICE DJD, 1970, COMMUNICATION SCI EN, P3	26
16	ALMIND TC, 1997, J DOC, V53, P404	25
17	CRONIN B, 2001, J INFORM SCI, V27, P1	25
18	ROUSSEAU R, 1997, CYBERMETRICS, V1, P1	25
19	SMALL H, 1974, SCI STUD, V4, P17	25
20	SCHUBERT A, 1986, SCIENTOMETRICS, V9, P281	24
21	COLE FJ, 1971, SCI PROGR, V11, P578	23
22	ZIPF GK, 1949, HUMAN BEHAVIOUR PRIN	23
23	BORGMAN CL, 1990, SCHOLARLY COMMUNICAT, P10	22
24	BORGMAN CL, 2002, ANNU REV INFORM SCI, V36, P3	22
25	MERTON RK, 1973, SOCIOLOGY SCI THEORE, P342	21
26	GARFIELD E, 1955, SCIENCE, V122, P108	20
27	MACROBERTS MH, 1989, J AM SOC INFORM SCI, V40, P342	20
28	INGWERSEN P, 1998, J DOC, V54, P236	19
29	KESSLER MM, 1963, AM DOC, V14, P10	19
30	MARTIN BR, 1983, RES POLICY, V12, P61	19
31	BRAAM RR, 1991, J AM SOC INFORM SCI, V42, P252	18
32	CALLON M, 1986, MAPPING DYNAMICS SCI	18
33	MCCAIN KW, 1990, J AM SOC INFORM SCI, V41, P433	18
34	NICHOLAS D, 1978, LITERATURE BIBLIOMET	18
35	SMITH LC, 1981, LIBR TRENDS, V30, P83	18

Tabla 1. Documentos más consumidos en Estudios Métricos. Fuente: Web of Science, 2006. Datos trabajados en CiteSpace.

Se destaca en esta tabla la utilización del trabajo de Lotka (1926), seguido por los de Price, Pritchard, Bradford, Garfield, Small, White y Egge, entre otros. Sin embargo, es fundamental mencionar la presencia constante de Price (que aparece con cuatro documentos entre los más citados); de Garfield (con tres documentos); y de White, Borgman y Cronin (cada uno con dos documentos citados).

Aunque este estudio abarcó el periodo entre 1975 y 2006, aparecen como documentos más utilizados trabajos que datan de 1926 (Lotka), 1934 (Bradford), 1949 (Zipf), 1955 (Garfield), 1963 (Price y Kessler), 1969 (Pritchard), 1971 (Pegue), 1973 (Small y Merton), 1976 (Narin) y 1978 (Nicholas). Se refleja así la importancia de estos textos para la conformación teórica de los estudios bibliométricos y cienciométricos en el mundo.

Respecto a la indagación de la frecuencia, aparecen destacados algunos documentos contemporáneos como los de Cronin (2001); White (1989); Borgman (2002) y McCain (1990) que seguramente obtendrán, con el paso del tiempo, un mayor índice de citas por parte de la comunidad científica.

3.3 Aportaciones provenientes de redes de libros y páginas web

Para el tercer estudio de caso se empleó *TouchGraph* con el fin de averiguar, de una parte, cuáles son los libros publicados sobre Ciencia de la Información (*Information Science*) y, de otra, verificar qué hiperenlaces aparecen relacionados con la Universidad Carlos III de Madrid.

Para cada análisis se utilizaron recursos diferentes de *TouchGraph*. Para el estudio de los libros se empleó como herramienta la plataforma *Amazon*, que en la actualidad es la mayor compañía de comercio electrónico de libros. Mientras que para los *links* relacionados con la Universidad Carlos III de Madrid se utilizó Google, el mayor motor de búsqueda y recuperación de información en la *web*.

El resultado de la investigación sobre los libros se clasificó en ocho categorías, además del núcleo principal, que viene determinado por la disciplina:

- Núcleo – Ciencia de la Información (objeto de estudio)
 1. Aislamiento de documentación de Ficción (probablemente indizado con errores);
 2. Conjunto de temáticas sobre Data Mining, Recursos Digitales e Inteligencia Artificial;
 3. Nodo de documentos sobre Realidad Virtual y Nuevas Tecnologías;
 4. Libros sobre principios de las ciencias, con especial destaque a la Estadística, Matemática y Comunicación;
 5. Bloque con temas de Política y Desarrollo del Poder de la Información;
 6. Nodo central sobre programas de Tecnologías, Informaciones y Comunicaciones;
 7. Conjunto de libros que tratan de la Certificación Computacional;
 8. Nodo de contenidos que describen las Técnicas de la Biblioteconomía y de la Documentación.

Estos nodos (ocho conjuntos) están representados en la figura siguiente.

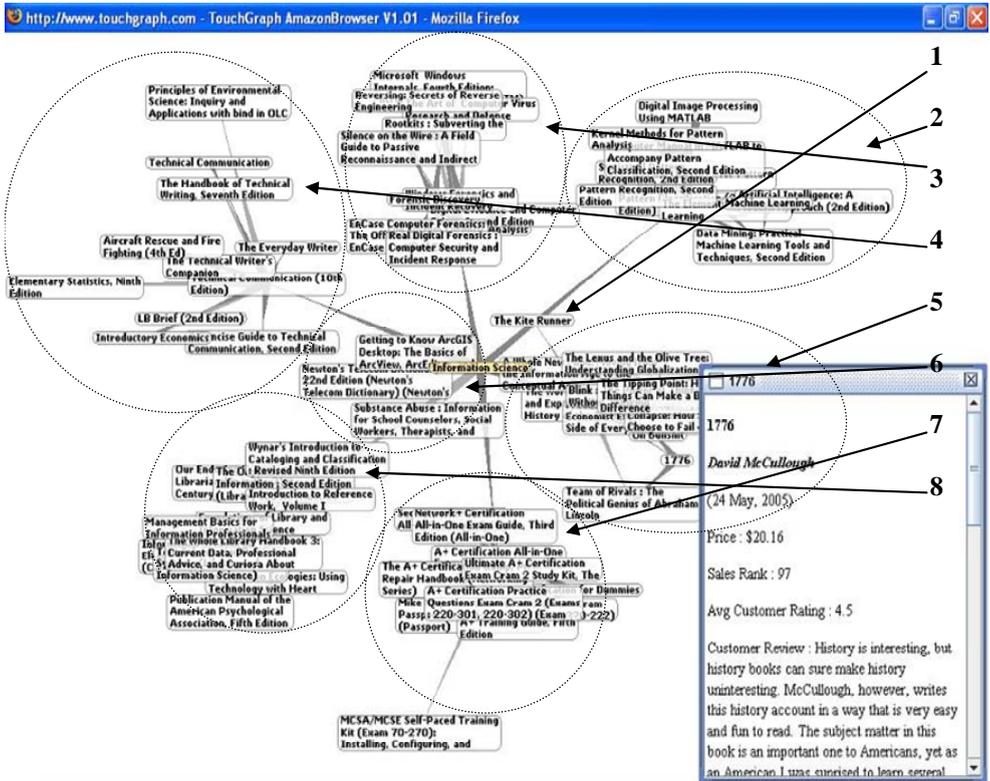


Figura 5. Red de libros sobre Ciencia de la Información en Amazon analizado con TouchGraph.

Es muy importante identificar las temáticas en que se distribuye la disciplina estudiada, ya que se trata de temas interrelacionados que se abordan dentro del ámbito ocupado por la Ciencia de la Información (Biblioteconomía, Documentación, Archivística, Informática, Estadística y CC. Políticas), cuyo mayor relieve se centralizó en torno a la Informática- Ciencia de la Computación, que agrega nuevas perspectivas científicas y que, a veces, parece que sea la ciencia mayor del grupo y no al contrario.

Una particularidad de este programa se deriva de las informaciones que ofrece sobre la identificación de los libros, su resumen y precio, favoreciendo una red de consumo pensada especialmente para la compra de productos editoriales.

Las relaciones representadas provienen del carácter exclusivo del programa, pues no se puede alterar el orden de ningún documento ni cambiarlo de posición.

La posición gráfica se deriva de la indización hecha por los profesionales de Amazon, que determinaron una Ciencia de la Información más fuerte en el ámbito informático, pero si la clasificación hubiese sido, con un mayor número de descriptores asignados al ámbito de la Información y Documentación la figura tendría el nodo 8 como el nodo central, junto a la expresión de búsqueda (*Information Science*).

El segundo análisis de este estudio de caso se hizo de una red universitaria que se aprecia interligada mediante Internet, más concretamente a través de enlaces con la Universidad Carlos III de Madrid, núcleo de la representación en este caso.

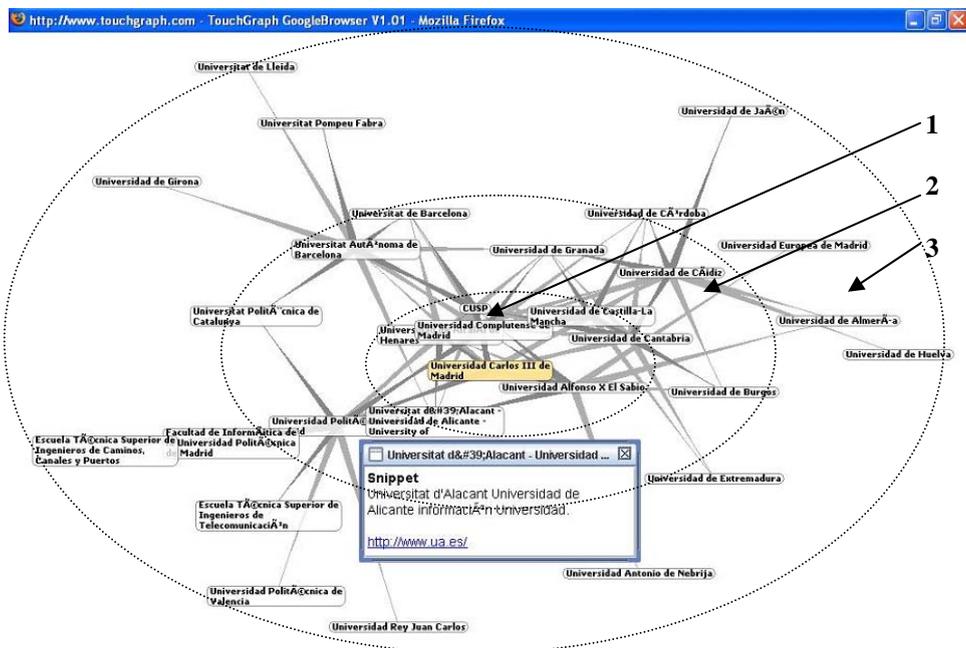


Figura 6. Contribución de universidades.

El análisis se dividió en tres categorías de relaciones, determinadas desde el vector de proximidad o desde el de distanciamiento, conforme se muestra en la figura 13.

Para efectuar el análisis, la herramienta *TouchGraph*, extendida a las investigaciones en Google, ofrece una nota con las informaciones básicas de cada relación. En este modelo identificamos informaciones provenientes de la Universidad de Alicante.

La agrupación de los nodos se hizo así:

- El primer grupo está constituido por la Universidad Alfonso X, la Universidad Complutense de Madrid, la Universidad Alcalá de Henares, el CUSP, la Universidad de Cantabria y la Universidad de Castilla-La Mancha.
- Para el segundo nodo las universidades representadas fueron las de Alicante, Autónoma de Barcelona, Barcelona, Burgos, Cádiz, Granada, Extremadura, Politécnica de Madrid y Politécnica de Cataluña.
- El grupo de menor centralidad está constituido por las Universidades Antonio de Nebrija, Rey Juan Carlos, Politécnica de Valencia, Girona, Pompeu i Fabra, Lleida, Jaén, Arcahueja, Europea de Madrid, de Almería y de Huelva, entre otras.

Otras relaciones obtenidas se dan entre las Universidades de Granada y la de Extremadura; la Universidad de Cádiz con la de Burgos; la Universidad de Cantabria con la de Cádiz; y de la Universidad de Barcelona con las universidades Pompeu i Fabra y Politéc-

nica de Cataluña. Fue posible visualizar estas relaciones gracias a que *TouchGraph* no sólo genera enlaces existentes con la dirección *web* investigada, sino que también establece relaciones entre los resultados que, en ocasiones, son más intensos que con el nodo central, la Universidad Carlos III de Madrid.

Este programa ofrece otra forma de visualización que consiste en asignar una tonalidad más intensa a las cooperaciones más fuertes, y tonalidades claras a las que tienen un intercambio menor. En este aspecto es importante destacar la constante relación de links de la universidad estudiada con el CUSP, la Universidad Complutense de Madrid, la Universidad Politécnica de Madrid y la Universidad Alfonso X, que podremos afirmar que son los principales aliados en las investigaciones científicas, según nuestra muestra.

4 CONSIDERACIONES FINALES

En la consolidación de las Redes Sociales son de vital importancia las relaciones. Por ello, sólo se puede estudiar una sociedad (grupos de personas, raza, grupos científicos,...) cuando se la conoce, puesto que, en caso contrario, la red puede hacerse con aportaciones que realmente no representen su estado correcto.

Para visualizar correctamente los resultados (mapas de relaciones) debe de tenerse en cuenta que los nodos de mayor cooperación, en general, están representados con una tonalidad más oscura, así como que la centralidad de algunos autores está representada como el punto del que parte la constitución del resto de relaciones.

Puede añadirse que los programas empleados como herramientas de aplicación en este estudio son complementarios, ya que su utilización como soporte en uno de los casos agrega informaciones a los demás y así recíprocamente entre los tres programas utilizados.

El debate sobre los estudios de caso no es fundamental, mientras que son de vital importancia las técnicas de análisis de los mapas.

Respecto a *Pajek*, debe de destacarse que trabaja desde una visión cartesiana manual, lo que supone una ventaja en el sentido de que no necesita extraer datos de otra plataforma, y su desarrollo tiene que irse creando paso a paso, conforme las necesidades del estudio. Otra ventaja es la posibilidad de realizar cálculos de densidad, cercanía, grafo, intermediación, clusters y vectores.

En relación a *CiteSpace* su visualización es muy agradable, pero su desventaja es que los datos primarios, sobre los que se generan los mapas, deben atender a los moldes de las plataformas de la *Web of Science* y de *PubMed*. Su mayor ventaja reside en la configuración gráfica automática de frecuencias y centralidades, lo que conlleva una mayor facilidad de estudio en los análisis de Redes Sociales.

Finalmente, el análisis de *TouchGraph* agrega una nueva vertiente a las Redes Sociales, como es averiguar las cooperaciones y relaciones que se dan mediante la *web*. Esta plataforma utiliza motores de búsqueda de tercera generación y plataformas de comercialización electrónica. Su principal ventaja está en ser una herramienta libre, que ofrece informaciones del entorno investigado, que muestra relaciones directas así como las de intensidad menor (representadas por la tonalidad de los hilos) y en que se puede clasificar su contenido en varios nodos, según su proximidad y centralidad. Sin embargo, es un recurso en el que no se pueden realizar relaciones de páginas que no aparezcan indizadas en *Google* y en *Amazon*.

5 BIBLIOGRAFÍA

BARNES, John A. (1954), Class and committee in a Norwegian Island Parish. *Humans Relations*, London, nº 7, p. 39-58.

BATAGELJ, Vladimir et Andrej MRVAR (2003). Pajek analysis and visualization of large networks. *Preprint Series*, Ljubljana, vol. 41, nº 871, p. 2-26.

BOURDIEU, Pierre (1977). *Outline of a theory of practice*. New York: Cambridge University.

CARTWRIGHT, Dorwin et HARARY, Frank (1956). Structural balance: a generalization of Heider's theory. *Psychological Review*, Washington, vol. 63 nº 5, p. 277-293.

CHEN, Chaomei (2006), CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, Maryland, vol. 57, nº 3, p. 359-377.

COLEMAN, James S. (1988). Social capital in the creation of human-capital. *American Journal of Sociology*, Chicago, vol. 94 supplement, p. 95-120.

CRANE, Diana (1972). *Invisible colleges: diffusion of knowledge in scientific communities*. Chicago, University of Chicago.

FISCHER, G. (1959). Business-Management social capital. *Zeitschrift fur Betriebswirtschaft*, Wiesbaden, vol. 29, nº 12, p. 725-736.

GLUCKMAN, Max H. (1954). *Rituals of rebellion in South-East Africa*. Manchester, Manchester University.

GRANNOVETTER, Mark S. (1985). Economic action and social structure: the problem of embeddedness. *American Journal of Sociology*, Chicago, vol. 91 nº 3, p. 481-510.

HEIDER, Fritz (1946). Attitudes and cognitive organization. *Journal Psychology*, Washington, vol. 21, nº 2, p. 107-112.

LITWIN, Howard (1997). Support network type and health service utilization. *Research on aging*, London, vol. 19, nº 3, p. 274-299.

LOMAS, Robert (2006). *El colegio invisible: la Royal Society, la francmasonería, el nacimiento de la ciencia moderna y la era de la razón*. Madrid: Martínez Roca.

MOLINA, José Luis; MUÑOZ, Juan Manuel et DOMENECH, Miquel (2002). Redes de publicaciones científicas: un análisis de la estructura de coautorías. *Redes – Revista Hispánica para el análisis de Redes Sociales*, Barcelona, vol. 1. Retrieved June, 10, 2006, from <http://revista-redes.rediris.es/pdf-vol1/vol1_3.pdf>.

MORENO, Jacob Levy (1934). *Who shall survive?* New York: Beacon.

PONTON, Robert D. (1975). Birth of psychological novel: cultural capital, social capital and literary strategy at end of 19th-Century. *Actes de la Recherche en Sciences Sociales*, Paris, nº 4, p 66-81.

PUTNAM, Robert D. (1995). Bowling alone: America's declining social capital. *The Journal of Democracy*, Washington, vol. 6, nº 1, p. 65-78.

ROSE, A. (1961). Housing and social capital. *Canadian Journal of Economics & Political Science*, Toronto Ontario, vol. 27, nº 1, p. 123-125.

SCOTT, John (1991). *Social network analysis*. London: Sage.

Enlaces sobre programas libres de análisis de Redes Sociales

StOCNET <<http://stat.gamma.rug.nl/stocnet/>>

NetMiner <http://www.netminer.com/NetMiner/home_01.jsp>

MultiNet <<http://www.sfu.ca/%7Erichards/Multinet/Pages/multinet.htm>>
AGD <<http://www.ads.tuwien.ac.at/AGD/>>
Carter's Archive of S Routines <<http://erzuli.ss.uci.edu/R.stuff/>>
daVinci <<http://www.informatik.uni-bremen.de/uDrawGraph/en/index.html>>
Doug White's software for kinship network analysis
<<http://eclectic.ss.uci.edu/~drwhite/doug.html>>
DyNet <<http://www.atalab.com/index.php>>
EigToll <<http://web.comlab.ox.ac.uk/projects/pseudospectra/eigtool/>>
Gem3Ddraw <<http://www.info.uni-karlsruhe.de/~frick/gd/index.html>>
InFlow <<http://www.orgnet.com/>>
Visone <<http://visone.info/>>
yFiles <http://www.yworks.de/en/products_yfiles_about.htm>
UcINET <<http://www.analytictech.com/ucinet/ucinet.htm>>

Otros enlaces importantes

Reference.com <http://www.reference.com/browse/wiki/Social_network>
Review REDES <<http://revista-redes.rediris.es/>>
INSNA <<http://www.insna.org/>>
Social Networks Analysis of Tom Snijders <<http://stat.gamma.rug.nl/socnet.htm>>
Canal de Rede Social <<http://www.social.org.br/>>
European Social Network <<http://www.socialeurope.com/>>
ISNAE <<http://www.isnae.org/>>