

E.5. Fin de las listas: necesidad de bibliografías estructuradas

Por Carlos Benito Amat

Benito Amat, Carlos. "Fin de las listas: necesidad de bibliografías estructuradas". En: *Anuario ThinkEPI*, 2007, pp. 148-151.

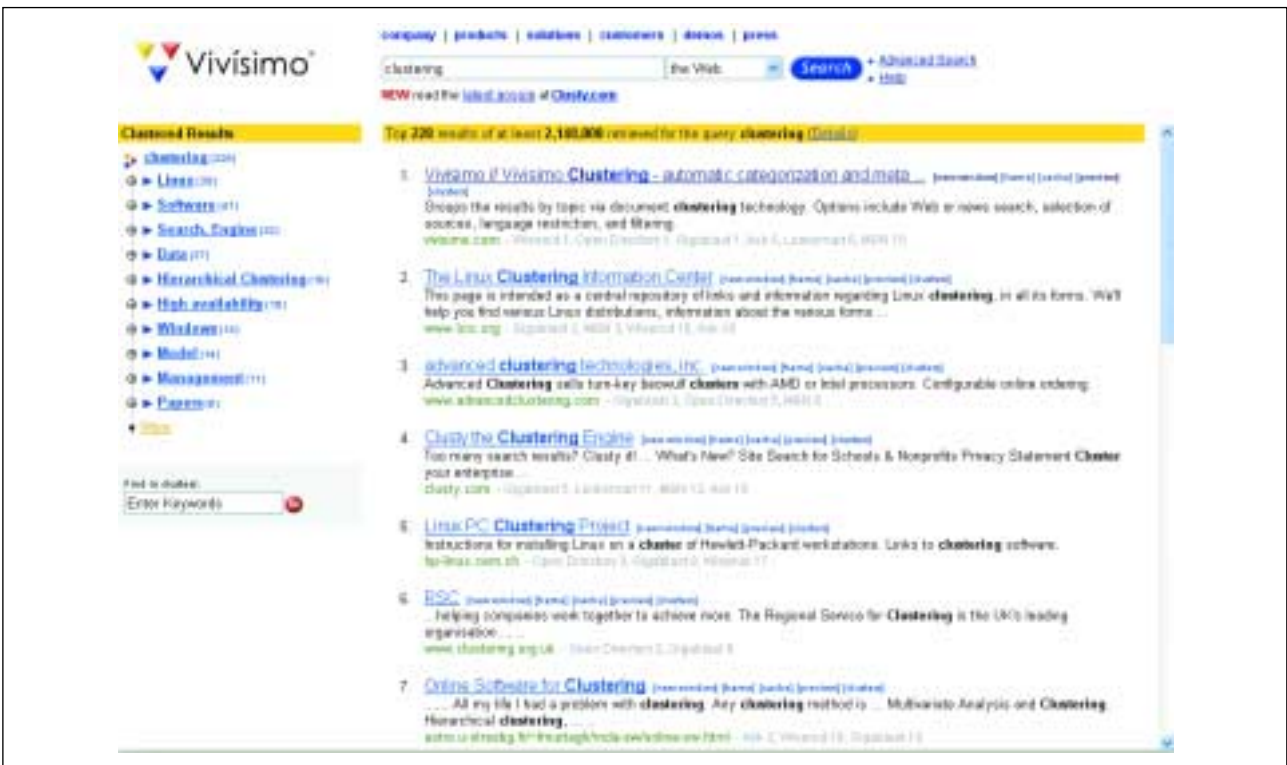


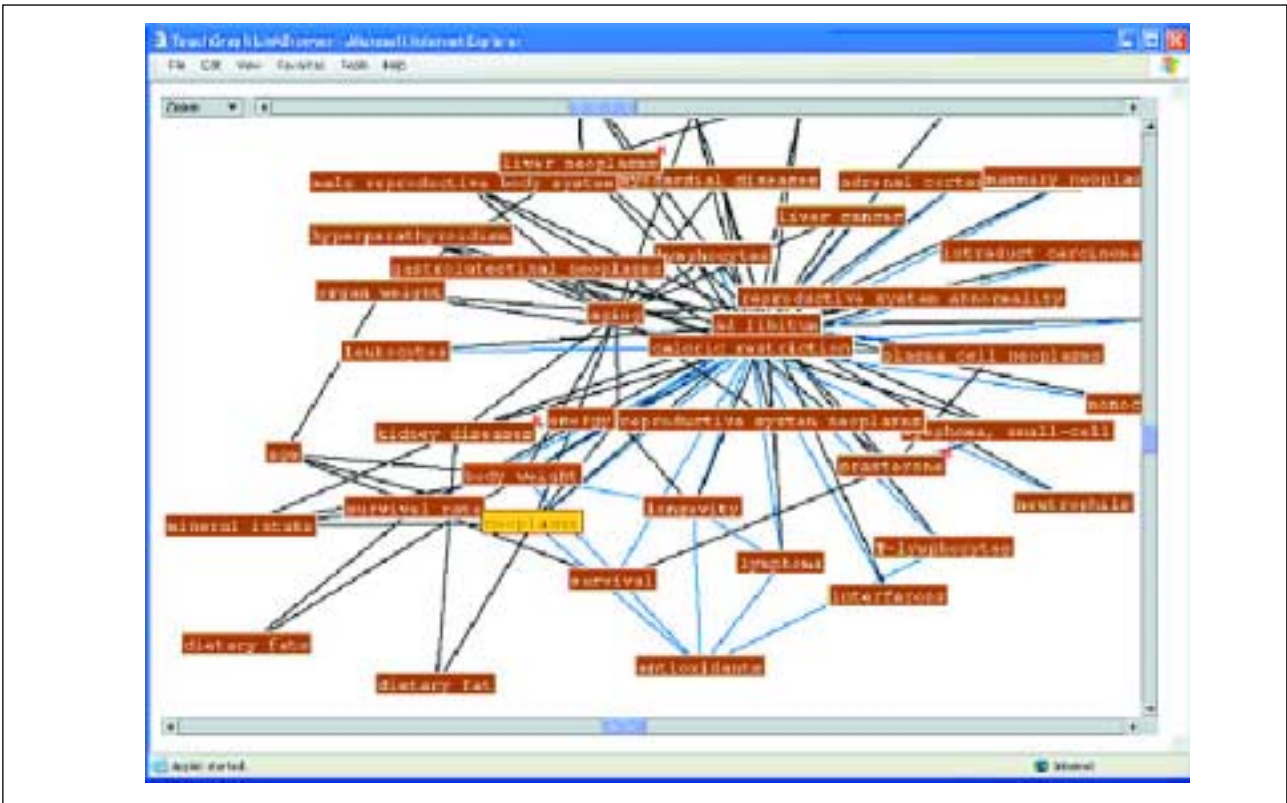
"Con una simple interconexión de recursos se podrían agrupar los resultados de consulta de varias bases de datos y catálogos según elementos comunes o relacionados"

Introducción

DURANTE SIGLOS las tareas de organización de la información han sido lineales, unidimensionales. Y la forma de distribuir esa misma información a los usuarios adoptaba la misma apariencia: los metros lineales de estanterías eran homólogos a las listas de referencias bibliográficas resultantes de cualquier consulta, manual o automática.

Hacia finales del siglo XIX, dos nuevos conceptos permitieron comenzar a mostrar alternativas a las listas informativas. Uno de ellos fue el sistema de archivo vertical, presentado en la Feria de Chicago de 1893 y cuyo descendiente más generalizado es el sistema de organización jerárquica de los archivos de ordenador¹. El otro procede del campo de la





información jurídica: las citas de precedentes legales que **Frank Shepard** transformó en producto en 1873² y que, casi 100 años después, inspiraron el desarrollo de los índices de citas científicas, además de continuar como referencia obligada en su propio campo hasta la actualidad. A los conceptos mencionados vino a añadirse un tercero, el de los esquemas clasificatorios facetados, implementados a partir de 1933 y cuyos conectores permiten poner en relación conceptos pertenecientes a jerarquías distintas³, situando un documento a lo largo de varias dimensiones.

Un contexto nuevo

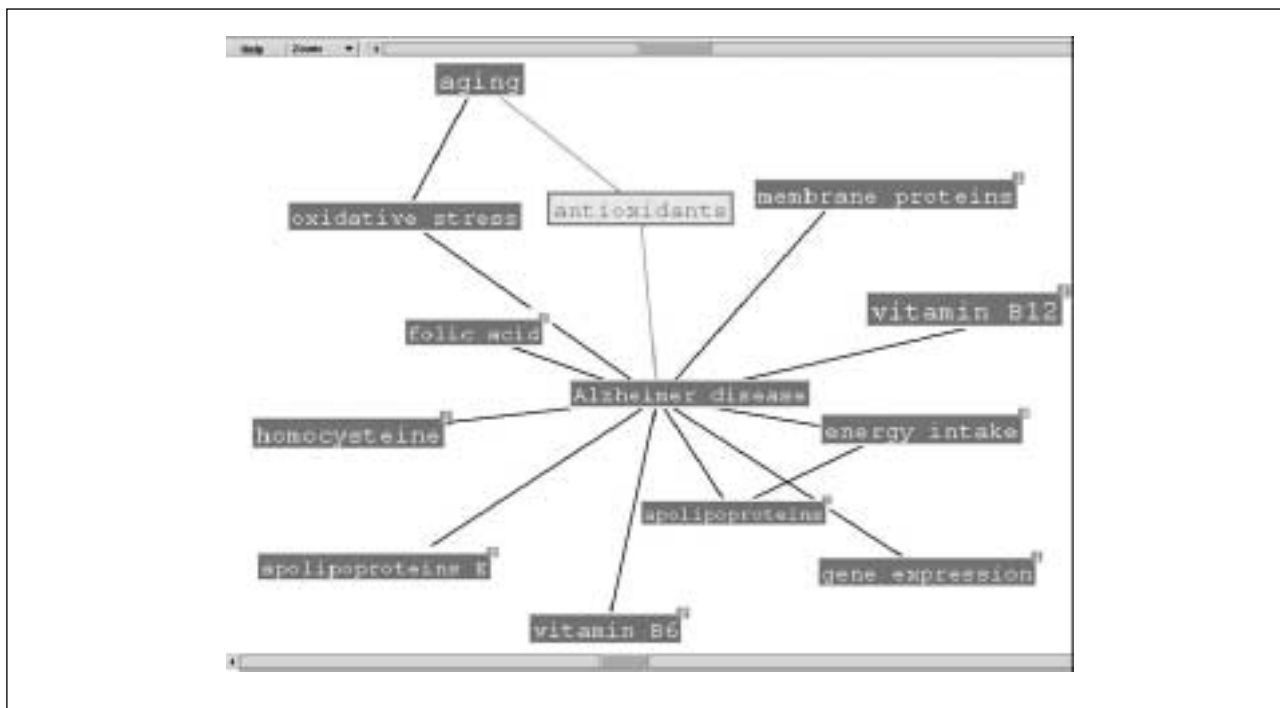
Los conceptos puestos en juego por la teoría de citas en documentación científica, especialmente el cálculo de influencias ponderadas de **Francis Narin** y **Gabriel Pinski**, figuran en el punto de partida del primer sistema de recuperación de información distribuida de tipo estructural⁴. No se trata de que un documento o una sede web reciba citas de otros documentos o enlaces a partir de otras sedes; importa, además, que esos documentos o sedes de partida sean significativas, importantes.

El sistema hipertextual de *Google* y sus seguidores configuran un nuevo contexto y obligan a reconsiderar la organización meramente lineal de la información y los documentos que la contienen. Y ello, entre otros motivos, por el carácter ingente, masivo del espacio informativo ahora accesible. Los usuarios se han venido acostumbrando a desplazarse a través del espacio Web para acceder a casi cualquier objeto informativo concebible.

Los principios básicos que definen este modelo de organización espacial de la información son dos:

1. Todo documento está relacionado con otro y esta relación se desprende de su estructura, y
2. la cuantificación de estas relaciones permite ordenar los documentos en espacios informativos.

Las afectaciones de una disposición legal, es decir, el conjunto de documentos anteriores que sustituye o enmienda y aquél otro de los documentos posteriores que la completan, corrigen o suprimen, es un ejemplo conocido del primer supuesto. Las listas de referencias bibliográficas de los documentos científicos y técnicos son otro.



Falta de ajuste

La mención de afectaciones y referencias en el párrafo anterior es deliberada. Las extraordinarias condiciones de consulta del *Boletín Oficial del Estado* quedan empañadas por la abundancia de sus resultados: una búsqueda genérica simple (*leche o productos lácteos*) ofrece más de 1.600 resultados, de los que 1.500 corresponden a documentos vigentes. No es la única base de datos que *muere a manos de su propio éxito*. La consulta de un concepto mucho más específico como “*determinación de flavonoides por cromatografía de alta resolución*” puede arrojar, según los recursos interrogados, entre 600 y más de 1.200 referencias; o entre 60 y 200 si se desean publicaciones recientes. Estos resultados sólo son abordables si quien demanda la información está dispuesto a costear su examen detallado. Casi nunca es el caso.

Y no sólo se trata de que los resultados de consultas sencillas en fuentes habituales sean inabarcables. También de la falta de ajuste entre el estado de conocimiento del usuario (desde lego hasta experto) y la dinámica de espacios informativos progresivamente masivos.

Iniciativas en marcha

Cualquier profesional de la información científica conoce *CrossRef*⁵; algunos han empleado los *citation* y *topic maps*, diseñados por *Inxight Software* para el sistema *HighWire*⁶, soluciones que transforman las relaciones entre documentos en hipervínculos, a veces dotados de expresión gráfica y que proporcionan un sentido de contexto en el recorrido de grandes bases de datos. Más allá de *Google*, se encuentran *Vivísimo*⁸ y su *clustered search* o la interfaz gráfica de resultados de *Mooter*⁹. Es muy reciente la adopción de *SFX*¹⁰ en los sistemas españoles de automatización de redes bibliotecarias. Muchos profesionales están expectantes en relación con *HistCite*¹¹, un sistema que idealmente ayudará al usuario a identificar con rapidez los trabajos más significativos sobre un tema¹², objetivo que lo sitúa por delante de los simples programas de visualización y diagramación de resultados de búsqueda.

Tras estas iniciativas se encuentran dos series de operaciones básicas:

1. El mero aprovechamiento de la interconexión de recursos, y
2. Procesamiento de las relaciones entre documentos para adjudicarles una posición determinada en espacios temáticos específicos.

Pérdida de oportunidades

Estas operaciones aún no se han puesto en marcha en nuestro entorno y los usuarios de la información están perdiendo interesantes oportunidades.

Así, una simple interconexión de recursos permitiría enlazar los resultados de una consulta en la base de datos de la *Agencia Española del ISBN*¹³ con los sistemas de las bibliotecas públicas o las redes académicas, a través del código ISBN. En muchos catálogos se podrían agrupar registros correspondientes a sucesivas ediciones de una misma obra con un ligero esfuerzo de procesamiento. Y los que consultan *IberLex*¹⁴ podrían obtener relaciones de documentos legislativos en cuya periferia se situarían las correcciones y enmiendas, y los textos que hubieran perdido su vigencia.

Al igual que en el caso de los índices de citas y de la búsqueda de documentos enlazados en el espacio Web, los usuarios tendrían también la oportunidad de juzgar la relevancia de los documentos relacionados si los resultados de cualquier búsqueda se expresaran de forma gráfica, situando los documentos más influyentes en el centro de una estructura que completaran aquellos otros interrelacionados con ellos. Hace casi 10 años que **Tim Bray**¹⁵ mostró una de las primeras formas de hacerlo. Se necesitan las bibliografías estructuradas: aquellas que muestren conjuntos de documentos interrelacionados y ordenados en un espacio temático definido.

Notas

1. **Yates, JoAnne.** "Business use of information and technology from 1880-1950". En: *A nation transformed by information: how information has shaped the United States from colonial times to the present*, eds. Alfred D. Chandler, Jr.; James Cortada, Oxford Press, 2000, pp. 107-135.
2. **DuCharme, Bob.** *A nineteenth-century linking application*. Permalink, 1 May 2003:
<http://www.oreillynet.com/pub/wlg/3153>
http://www.oreillynet.com/xml/blog/2003/05/a_nineteenthcentury_linking_ap.html
3. **Gil Urdiciain, Blanca.** *Manual de lenguajes documentales*. 2ª edición. Gijón: Trea, 2004, pp. 101-120.
4. **Page, Larry.** *Method for node ranking in a linked database*. US Patent 6.285.999. 9 Jan. 1998.
5. <http://www.crossref.org/>
6. <http://highwire.stanford.edu/>
7. <http://www.northernlight.com/>
8. <http://www.vivisimo.com/>
9. <http://www.mooter.com/>
10. <http://www.exlibrisgroup.com/sfx.htm>
11. <http://www.histcite.com/>
12. **Garfield, Eugen; Pudovkin, Al; Istomin, VS.** *Algorithmic citation-linked historiography: mapping the literature of science*. 65th ASIS&T annual meeting, Nov. 18-21, 2002. Philadelphia: ASIS&T.
13. <http://www.mcu.es/bases/spa/isbn/ISBN.html>
14. <http://www.boe.es/g/es/iberlex/>
15. **Bray, Tim.** *Measuring the Web. Computer networks and ISDN systems*, 1996, v. 28, n. 7-11, pp. 993-1005. Accesible en:
http://www5conf.inria.fr/fich_html/papers/P9/Overview.html