

DE LOS TESAURUS A LOS *TOPIC MAPS*: NUEVO ESTÁNDAR PARA LA REPRESENTACIÓN Y LA ORGANIZACIÓN DE LA INFORMACIÓN

OF THESAURI TO TOPIC MAPS: NEW STANDARD FOR THE REPRESENTATION AND THE ORGANIZATION OF THE INFORMATION

[José Moreiro González](#)

Juan Llorens Morillo

Miguel Ángel Marzal García-Quismond

Jorge Morato Lara

Pilar Beltrán Orenes

Sonia Sánchez Cuadrado¹

RESUMEN

El sustantivo crecimiento del número de tesauros no ha servido para responder a la necesidad de trabajar en entornos multidisciplinares. Para responder a esta situación, se fomentó, en un primer momento, la fusión de tesauros con objeto de adaptar los preexistentes a las necesidades planteadas por los nuevos dominios. Más tarde, en el afán de atender precisamente a un mundo de información cambiante y en crecimiento, se empezó a trabajar en los mapas conceptuales de navegación. Hasta desembocar en los *Topic maps* surgidos de la necesidad de fusionar índices para incorporar así la utilidad de los hiperenlaces. Su éxito se plasmó en la norma ISO/ICE 13250:1999, que motivó la hipótesis de su uso en la elaboración de tesauros, proceso en el que se manifestaron las contradicciones entre ambos sistemas. Por lo que se analizan los elementos conceptuales esenciales en la arquitectura de los *topics maps* para hallar su correlato en los tesauros. Este estudio comparativo permite señalar como límites de los *topics maps*, su indefinición conceptual, sus innumerables relaciones y su ambigüedad, lo que no oculta sus ventajas como una mayor riqueza semántica, la consecución de un nuevo marco conceptual para dominios fusionados, las asociaciones determinadas por verbos y la capacidad de organizar recursos informativos de distinto tipo.

PALABRAS-CLAVE: *Topic-Maps*. Tesauros. Mapas conceptuales. Representación de la información. Procesamiento automático de la información.

1 INTRODUCCIÓN

El sustantivo crecimiento del número de tesauros no ha servido para responder a la necesidad de trabajar en entornos multidisciplinares, por lo que en un primer momento se fusionaron tesauros con objeto de adaptar los preexistentes a las necesidades planteadas por los nuevos dominios². Más tarde, en el afán de atender precisamente a un mundo de información cambiante y en crecimiento, se empezó a trabajar en los conocidos mapas conceptuales de navegación. El estudio de los mapas temáticos se originó en la necesidad de crear índices en torno a algún concepto o materia.

Las redes semánticas son un método común de representar el conocimiento en el campo de la inteligencia artificial, que busca establecer comunicación entre las personas y las máquinas. Uno de los mecanismos más usados es el de los grafos conceptuales, cuyos elementos constituyentes son conceptos y relaciones de conceptos. Los grafos o redes son colecciones ordenadas de nodos conectados por arcos que se usan para representar documentos. Un ejemplo es el grafo denominado red semántica, que representa las relaciones semánticas que se establecen en el texto, y que se pierden (a menudo), en otros sistemas de indización. Estas redes semánticas o redes asociativas tienen su origen en las propuestas del filósofo inglés Charles Peirce en el siglo XIX que ideó notaciones gráficas para la lógica simbólica³. Aunque el precedente más conocido es el de los grafos teóricos propuesto por John Sowa⁴.

También los mapas conceptuales son una técnica para representar el conocimiento en gráficas. Las gráficas cognitivas son redes de conceptos. Las redes se componen de nodos (puntos / vértices) y enlaces (*arcs*: arcos / *edges*: márgenes, extremos, satélites). Los nodos representan conceptos y los enlaces representan las relaciones entre los conceptos. Comparados con el restringido número de relaciones de los tesauros, los mapas conceptuales ofrecen una red de relaciones más rica, sobre la que pueden soportar diferentes mecanismos de razonamiento. El concepto que origina su estructura (nodo-enlace-nodo) es muy próximo a su equivalente en las redes hipertextuales y por consiguiente soportan la navegación de un modo muy natural. El uso de mapas conceptuales permite desarrollar mejores mecanismos de representación y recuperación, ya que las relaciones entre los conceptos se eligen teniendo en cuenta la modelización de las necesidades y expectativas de cada usuario.

Índices, glosarios y tesauros son formas diferentes de esquematizar las estructuras del conocimiento. Podemos ver un grafo conceptual para una oración gramatical: “Juan pinta un cuadro”, donde los conceptos van entre corchetes, mientras las relaciones van entre paréntesis.

[Juan] <- (agente) <- [pintar] -> (objeto) -> [cuadro]

No fue hasta principios de los 90 cuando los Mapas Conceptuales se empezaron a utilizar. El Davenport Group buscaba medios que posibilitasen el intercambio de documentos entre ordenadores, e idearon el *DocBook*⁵, DTD muy conocido para desarrollar documentos SGML y XML.

Los conceptos y, a veces, los enlaces están etiquetados. Los enlaces pueden ser no direccionales, unidireccionales o bidireccionales. Los conceptos y los enlaces pueden categorizarse, pudiendo ser solamente asociativos, especificados o divididos en categorías como relaciones causales o temporales.

Los mapas de conceptos pueden hacerse con varios fines⁶:

- Para generar ideas (tormenta de ideas, etc.);
- Para diseñar estructuras complejas (textos largos, hypermedia, sitios web, etc.);
- Para comunicar ideas complejas;
- Para ayudar al aprendizaje, al hacer patente la integración del nuevo conocimiento con el viejo.

La construcción de un mapa conceptual sigue este proceso⁷:

- 1) Selección de los conceptos que se representarán en el mapa.
- 2) Listado de esos conceptos.
- 3) Agrupación de los conceptos relacionados.
- 4) Ordenación de los mismos en forma bidimensional o tridimensional.
- 5) Enlace de cada par de conceptos mediante líneas etiquetadas en modo preposicional o proposicional.

El mapa conceptual viene a complementar la función comunicativa del conocimiento que incumbe al lenguaje natural, al utilizar diagramas confeccionados mediante diferentes lenguajes visuales para representar gráficamente los conceptos y las relaciones que se dan entre ellos. Son una herramienta propicia a la hora de organizar nueva información e integrarla en el conocimiento existente, ya que su construcción ayuda a reconocer nuevas relaciones entre conceptos y a refinar la comprensión de las relaciones existentes.

2 LOS TOPIC MAPS

Un paso más, dentro de los mapas conceptuales, se daría al aparecer el nuevo estándar ISO/ICE⁸. Se partía del convencimiento de que en cualquier actuación que buscarse establecer un “*principio fundamental de organización para la creación y mantenimiento de la información*”⁹, tenía que aprovecharse obligatoriamente la utilidad que ofrecen los hiperenlaces. El fin de los *Topic Maps* consiste en proporcionar acceso a la información existente en diferentes redes semánticas, como pueden ser los propios tesauros. De hecho, su aplicación inmediata han sido los diferentes recursos que pretenden proveer el acceso a la información desde algún modelo de conocimiento, así como la generación de nuevos recursos

terminológicos: equivalentes electrónicos en índices, tablas de contenidos, glosarios, tesauros, referencias cruzadas, etc. No debe extrañarnos que se utilicen los *Topic Maps* para crear tesauros, puesto que los tesauros no son si no una simplificación del esquema, más amplio, que proponen aquellos. Veamos primero las características del estándar *Topic Maps*, tanto en su naturaleza instrumental como conceptual, para luego señalar las diferencias y semejanzas que existen entre esta propuesta y la más clásica y conocida de los tesauros.

La Norma ISO/IEC 13250 *Topic Maps* define el modelo y la sintaxis de intercambio para formalizar Mapas conceptuales. Esta norma internacional aporta los signos estandarizados necesarios para representar la información de los diferentes recursos desde la definición de los conceptos y las relaciones que existen entre ellos¹⁰. Así, un conjunto de uno o más documentos que utiliza la notación definida por este estándar origina un mapa conceptual o temático (*topic map*). Precisamente por la estructuración semántica que ofrece de los enlaces en la red se denomina a esta norma “el GPS del universo de la información”, en cuanto solución para navegar y organizar los extensos y continuamente crecientes recursos de información (publicaciones de todo tipo y, en especial, sitios web), y para establecer un puente entre los campos del conocimiento y la gestión de la información.

La puesta en práctica de los TOPIC MAPS se ideó para arquitectura SGML con notación de HyTime, la DTD del estándar ISO/IEC 13250:2000 se ha mantenido en HyTime. No obstante, esta notación cayó en desuso tras la aparición de XML. En este caso, las causas del desplazamiento de la notación de HyTime se deben a la creación de una DTD para crear Topic Maps en XML, denominada X TOPIC MAPS, y al auge que ha obtenido XML. En cualquier caso, actualmente se pueden encontrar ejemplos de ambos lenguajes en la Web¹¹. Dada la novedad del estándar TOPIC MAPS y su, todavía, escasa implantación en la web, se ha considerado oportuno revisar, en primer lugar, sus elementos principales, para luego centrarnos en su análisis y perspectivas futuras.

Las ideas fundamentales de la norma son la representación de los conceptos o elementos de recuperación de la información, los diferentes casos que ofrecen (agrupamientos de objetos direccionables de información alrededor de los conceptos), y las relaciones (asociaciones) que se dan entre ellos. Los conceptos clave que conforman un *topic map* son¹²:

- *Topic*
- *Topic Type*
- *Topic Occurrence*

- *Occurrence Role*
- *Topic Association*
- *Association Type*
- *Scope y Theme*
- *Public Subject*
- *Facets*

En sentido amplio, un *topic* representa en el *topic map* a una persona, entidad o asunto, es decir, a algún concepto abstraído de una realidad cualquiera. La abstracción puede ser individual, es decir, puede referirse a sujetos particulares (por ejemplo, Madrid, que sería la abstracción de la ciudad existente en la realidad), o puede hacer referencia a sujetos generales (por ejemplo, ciudad, que se correspondería con la abstracción de todas las ciudades que han existido en la historia de la Humanidad). Formalmente, el término concepto (*topic*) designa un elemento perteneciente a un mapa conceptual que representa a una materia (*subject*), esta existe solo como referente (ideal) de los *topic*, pues para denominarlo sería necesaria una definición. En cierta manera, como representante abstracto de un concepto, una materia (*subject*) se corresponde con lo que Platón llamaba *idea* en el mito de la caverna: los términos son sombras de los conceptos, algo que no se puede representar realmente. Así, varios *topics* pueden referirse al mismo *subject*, ya que, entre otras posibilidades, caben los sinónimos. Un *topic* es un concepto infinito que, consecuentemente, nunca puede ser atrapado en su totalidad. Un tipo de concepto, en fin, que nunca queda definido totalmente porque depende de la casuística que viene dada por los *topics*. Sin embargo, habitualmente ambos significados se fusionan cuando nos referimos a ellos usando el término concepto.

Los conceptos (y sus respectivos tipos) varían según varíe la información que se esté tratando. En efecto, si se está hablando de documentación relativa a *software*, los conceptos representarán funciones, variables, objetos y métodos informáticos, mientras que si la información es relativa a botánica, los tópicos se referirán a plantas, algas y sus funciones¹³. Partiendo de esta definición de *subject*, y teniendo presente en todo momento que el *topic* es la representación en lenguaje XML del *subject*, consideremos ahora los conceptos clave que conforman un *topic map*:

Los *topics* tienen tres características principales: su denominación (*Topic names*), sus apariciones o casos (*occurrences*) y su rol en las asociaciones (*role associations*). Dos *topics* con características equivalentes se consideran semánticamente idénticos.

El *topic name* hace referencia a las diferentes formas de denominación que puede tener un *topic*: así León, capital leonesa, Legio VII, o Le- son diferentes formas de llamar al mismo concepto. Existen varios tipos de nombres posibles para un *topic*: formal, simbólico, apodo, ... La forma que el estándar propone para el nombre incluye obligatoriamente la forma normalizada (*base name*). El *base name* es un elemento obligatorio y representa la forma usual de hacer mención al *topic*. Es complementario del identificador interno (id). Los *base name* deben de ser únicos en un dominio determinado bajo un *scope* dado. Además el *topic* puede tener otras denominaciones (*alternative names*), como el (*display name*), que es la forma en la que se mostrará al usuario, y el (*sort name*), que es como se ordenará alfabéticamente cuando se saque un listado.

Mientras que el *topic type* es una manera de agrupar los conceptos, mediante la relación que se establece entre una clase y sus instancias, lo que guarda un gran parecido con la relación clase-instancia jerárquica de los tesauros. Un *topic type* podría ser “país” y sus instancias los *topics* “España”, “México”, “Brasil”, “Colombia”, etc. Aunque los *topics* sean considerados como instancias de los *topic types*, éstos a su vez pueden ser considerados como *topics*. Así, un *topic* puede ser “Madrid” y también puede serlo “ciudad”, sin que esto impida que, al mismo tiempo, “Madrid” sea una instancia del *topic type* “ciudad”. Los *topic types* varían en función del tipo de información con la que estemos tratando, es decir, los conceptos que dan lugar a los *topics* cambian en cada caso.

Cada *topic* tiene, al menos, una aparición o caso (*occurrence*), que es un puntero dirigido a la fuente de información donde el concepto es relevante. Las ocurrencias pueden ser de varios tipos; cada uno de ellos se denomina *occurrence rol*. Las *occurrences* son los casos relevantes, o ejemplos significativos de un *topic* y suelen estar fuera del documento, o documentos, que se toma como referente del *topic map* en cuestión. En sentido estricto son los *links* a otros recursos informativos en los que aparecen los ejemplos citados, y que, aunque ajenos al *topic map*, mejoran la comprensión de un *topic*. Así, los *topics* y sus *occurrences* se sitúan a dos niveles distintos. Igual que en el caso de los *names*, un *topic* puede estar enlazado (*link*) a una o más ocurrencias. Al igual que los *topics* se pueden clasificar en *topic types*, las *occurrences* se pueden agrupar en *occurrences role*. Las *occurrences* pueden ser de distintos tipos en función del recurso al que se enlace, esto es, pueden ser páginas *web*, artículos, monografías, comentarios, etc., y a esto es a lo que se le conoce como *occurrence role*. Debe resaltarse que la mayoría de estas *occurrences* suelen ser externas al TOPIC MAPS, siendo una situación análoga a la que siempre ha existido entre un tesoro y la indización de documentos que se realiza con su aplicación.

Contamos con dos tipos de *ocurrences*: *resourceRef* que es un enlace a un recurso externo de información; y *resourceData* que es algún dato no externo que se facilita, como puede ser una definición.

Las ideas de *topic*, *topic type*, *topic name*, *ocurrence* y *ocurrence rol* permiten organizar las fuentes de información de acuerdo con los *topics*, así como establecer índices.

El *topic* no se define solo por su denominación (*topic name*), si no por sus relaciones (*associations*) y su ámbito (*scope*). En un tesauro se define el término mesa, pero en los *topic maps* se define *esta* mesa, que *está en* este edificio, que *sirve* para esto, ... de forma que se puede trabajar con instancias, con elementos concretos, mientras que el tesauro se inclina hacia lo genérico de los términos. Precisamente, esta norma alcanza su mayor importancia por la riqueza de agrupaciones que permite establecer *asociaciones* entre los conceptos, de acuerdo con las posibilidades que ofrece la lógica y la semántica de mundo, la vida. Cada concepto participante en una asociación posee un *rol de asociación* que determina el papel que desempeña el concepto en la asociación. La *topic association* es la relación que se establece entre *topics*. Formalmente es un elemento que enuncia la relación entre dos o más *topics*. Las *topics association* están compuestas por dos o más *topics* y por las formas verbales que los une. Un ejemplo puede ser:

Diego Velázquez PINTÓ Las Meninas

[<Topic><Association><Topic>],

donde el concepto *Diego Velázquez* cumple el papel de *Pintor*, que a su vez sería otro concepto (*Topic tipe*), mientras que a *Las Meninas* le correspondería el papel de *cuadro*, que a su vez es un nuevo concepto (*Topic tipe*). Los tipos de asociación (*Association tipe*) engloban cuantas asociaciones puedan darse: sinonimia, ubicación, generalización, vivir en, ... Si bien hay que determinar el ámbito en que esto es cierto, que Diego Velázquez sea un pintor español del siglo XVII se determina por el *Theme Scope*, el tema o ámbito en que una relación es cierta. De esta manera podemos considerar que las asociaciones pueden ser transitivas, además pueden tener dirección, de forma que sean simétricas: si A conoce a B, entonces B conoce a A, o asimétricas como sería el caso: si A nació en B, no puedo decir que B nació en A. Las asociaciones también pueden ser transitivas, es decir que si A es igual a B y B es igual a C, entonces se puede afirmar que A es igual a C. Esta cualidad ha permitido un desarrollo muy interesante alcanzándose la definición de *asociación taxonómica*, dentro de la que distinguen dos categorías:

1. **Tipos de asociaciones básicas**, una selección de las relaciones existentes en los tesauros, con prioridad en la relación parte – todo.
2. **Asociación de propiedades**, a partir de relaciones binarias y empleando como elementos de definición las anteriormente enunciadas propiedades algebraicas. De este modo se clasifican cuatro asociaciones según la combinación de propiedades:
 - *Relación de equivalencia*, aquella en la que los *topics* cumplen las propiedades reflexiva, simétrica y transitiva.
 - *Relación de comparación parcial*, donde el *topic* cumple las propiedades reflexiva, antisimétrica y transitiva.
 - *Relación de comparación total*, que concede al *topic* las propiedades reflexiva, antisimétrica y transitiva, pero implementadas con una propiedad conectiva.
 - *Relación de comparación “más fuerte”*, caso que reconoce al *topic* las propiedades antireflexiva, antisimétrica y transitiva.

Como en los casos anteriores, las *topic associations* se pueden clasificar según la *association type*. La *association type* está definida por la forma verbal que une a los *topics*, es decir, la *association type* es la que define el verbo que une en cada caso a los *topics*. Ejemplos pueden ser: “nació en”, “escrito por”, “pertenecer a”. Todos estos casos distintos constituyen, a su vez, los diferentes tipos de asociación. Cada una de las *association types* puede ser interpretado también como un *topic*. *Association role* es la función desempeñada por cada uno de los *topics* que participa en una asociación concreta. Por ejemplo, en “nació en” podríamos decir que los dos *topics* que une, deben ser la *association role* <personaje>/<lugar>. Cada vez que se asigna una característica a un *topic*, bien sea un *name*, una *occurrence* o un *role*, sólo se considera válida dentro de ciertos límites o contexto. Estos límites de validez pueden estar especificados de manera más o menos explícita y constituyen el *scope*. El límite de la validez de las asignaciones de cada *topic* se expresa mediante el conjunto de temas (*theme*) en los que las asignaciones tengan lugar. Un ejemplo podría ser “banco” en el tema “financiero” o en el tema “pesca”.

Algunas veces el mismo constructo es representado por más de un *topic link*. Esto ocurre cuando se mezclan dos *topic maps* (por ejemplo, “France” y “Francia” pueden ser dos *topics* distintos que en realidad representan el mismo *subject*). En estos casos hay que tener alguna forma para establecer la identidad entre *topics* aparentemente diferentes. El concepto que posibilita esta identidad es el de *public subject*, y el mecanismo usado es un atributo (el

identity attribute) en el elemento *topic*. Estos atributos se dirigen a un recurso o fuente electrónico que identifica al *subject* en cuestión de la manera menos ambigua posible. Esta fuente puede ser oficial, un documento validable públicamente (por ejemplo una norma ISO) o puede ser simplemente una definición descriptiva dentro (o fuera) de uno de los *topic maps* considerados.

Cualquiera dos *topics* que se refieran al mismo *subject* por significado de sus *identity attributes* son considerados semánticamente equivalentes a un *topic simple* que reúne las características de ambos *topics* (los nombres, ocurrencias y asociaciones).

Los *Topic Maps* pueden dividirse, como los tesauros, en facetas (*facets*), con la diferencia de que en este caso las facetas pueden afectar también a las ocurrencias, es decir, a los documentos con determinada información. Así, se puede pedir una faceta limitada tan solo a recuperaciones que actúan como filtros con informaciones, por ejemplo, en páginas web, o en enciclopedias, o en documentos escritos en portugués. Las facetas deben ser consideradas como atributo-valor, en el sentido de que son propiedades. Las *facets* proporcionan un mecanismo para asignar pares de property - value de recursos o fuentes de información. Una *facet* es simplemente una propiedad, y sus valores son denominados *facet values*. Las *facets* son usadas normalmente para suplir la clase de metadatos que pueden ser provistos por atributos SGML o XML. Pueden incluir propiedades como “lenguaje”, “seguridad”, “aplicabilidad”, etc. Las *facets* pueden ser usadas también para cubrir las clases de propiedades usadas en los sistemas de clasificación facetada. Por último, las *facets* constituyen un complemento al *scope*, mientras que el *scope* puede ser visto como un mecanismo de filtrado que está basado en las propiedades de los *topics*, las *facets* proporcionan un filtro basado en las propiedades de las fuentes de información mismas. En este sentido, las *facets* son ortogonales al modelo del *topic maps* mismo (excepto cuando *facets* y *facet values* son vistos como *topics*).

La notación base de los Mapas conceptuales se establece en SGML. Un mapa conceptual contiene al menos un documento SGML, si bien puede incluir otros tipos de recursos informativos o referirse a ellos¹⁴. Un grupo de recursos de información que comprende un completo mapa de conceptos intercambiable puede especificarse utilizando la ayuda del grupo de objetos enlazados *Bounded Object Set* (BOS)¹⁵. Dado que XML es un subgrupo de SGML, puede ser también usado como notación base para los mapas conceptuales¹⁶. Así pues, un mapa conceptual en su formato de intercambio es un documento o grupo de documentos SGML (o XML) interrelacionados que emplea la notación definida por esta Norma¹⁷. En su interior se define un espacio donde las localizaciones son conceptos,

y donde las distancias entre conceptos son apreciables a partir del número de conceptos intermedios que deben ser visitados si se quiere ir desde un concepto a otro, así como el tipo de relación que define la trayectoria de un concepto a otro, si la hay, a través de los conceptos intermediarios, si estos existen¹⁸. La arquitectura de los mapas conceptuales está ideada para facilitar la unión de mapas conceptuales sin la necesidad de combinar otros mapas temáticos para copiarlos o modificarlos.

Esta norma alcanza su función más extensa en su posible aplicación a los recursos de información comerciales y corporativos, no limitándose tan solo a aquellos publicados en papel. Para todos ellos supone un valor añadido, al servir aquella información que los usuarios realmente necesitan. La utilidad de su ayuda en línea conlleva facilitar el acceso a la masa informativa. En especial para los portales de información, ya que pueden organizar los sitios web mediante mapas, a la vez que ofrecer enlaces metainformativos con otros sitios web, al combinar adecuadamente los módulos de datos, con lo que se logra simplificar el acceso a la información relevante.

Surge de manera inmediata su aplicación a la elaboración de obras de referencia como los diccionarios y las enciclopedias, así como de documentación técnica. El problema al que se enfrentan los diccionarios o las enciclopedias es que la mayor parte de la información que contienen se puede encontrar en algún sitio al que se accede libremente mediante la red. Entonces, si quieren competir con ventaja, deben proporcionar información de calidad y hacerlo de la manera más rápida posible. Labores ambas a cuya excelencia ayudan los mapas de temas. Por su parte, la documentación técnica, que suele componer cientos de páginas con datos sobre los productos que describen, exige una actualización constante causada en las sucesivas renovaciones del producto. Esas continuas versiones requieren una organización de la información que va más allá de un libro corriente.

Pero, desde luego, los mapas conceptuales son candidatos ideales para aplicarlos a cualquier tipo de texto y, en especial, a los científicos. Modular el texto es el primer paso para llegar a una solución adecuada¹⁹. La superestructura específica (capítulo-sección) de un texto de documentación científica se fracciona en cientos de microestructuras, algunas de las cuales se establecen como autocontenidos que se muestran divergentes respecto al tema concreto de la macroestructura. Los mapas conceptuales conceden una perspectiva múltiple y concurrente de colecciones de documentos. La naturaleza estructural de estos enfoques es libre; pueden reflejar una aproximación orientada a objetos, o pueden ser relacionales, jerárquicas, ordenadas, desordenadas o mostrar cualquier otra posible combinación. Lo cual nos permite utilizar los mapas conceptuales:

- Para disponer como conceptos el contenido y los datos registrados en los documentos y permitir herramientas de navegación tales como índices, referencias cruzadas, sistemas de citas, o glosarios.
- Para agrupar conceptos enlazados y permitir la navegación entre ellos. Esta capacidad puede usarse para el ensamblaje de documentos virtuales, y para generar tesauros en cuanto interfaces hacia los corpora, las bases de conocimiento, etc.
- Para tamizar conjuntos de informaciones y establecer perspectivas adecuadas a usuarios o propósitos específicos. Así, este filtrado puede ayudar en la gestión de documentos multilingües, gestionar los modos de acceso dependiendo de los criterios de seguridad, suministrar visiones parciales de acuerdo con el perfil del usuario o los dominios del conocimiento, etc.
- Para estructurar los objetos de información desestructurados o para facilitar la creación de interfaces de usuario, orientadas a conceptos, que posibilitan la fusión de bases de información desestructuradas con otras estructuradas. El mecanismo de incrustación de los mapas conceptuales puede ser considerado como una especie de "mecanismo externo de valor añadido", en el sentido de que se impone una estructura arbitraria a la información sin alterar su forma original.

Respecto a la visualización, los *Topic Maps* aportan una de las principales propuestas para visualizar la web semántica²⁰. El hecho de que un *Topic Map* puede presentar miles de asociaciones de diferente tipología (*association types, roles, occurrences, etc*) supone un problema. Actualmente, entre las representaciones propuestas destacan los árboles, los *browsers* y los gráficos. A la hora de analizar estas representaciones se debe tener en cuenta que muestren tanto información local de los *topics* que interesan al usuario, como información sobre la localización de esos *topics* en el conjunto del *Topic Map*. Usar para ello árboles con hipervínculos es uno de los principales inconvenientes de este planteamiento, pues si bien la percepción del *Topic Map* es más sencilla para el usuario, no pasa lo mismo con los distintos elementos del *Topic Map*. La tipología de los *Topic Maps* (*topics, occurrences, roles, types*) hace que el usuario pueda desorientarse ante un árbol de estas características.

3 RELACIÓN DE LOS *TOPIC MAPS* CON OTROS RECURSOS QUE REPRESENTAN EL CONOCIMIENTO

Los *Topic Maps* se proponen para modelar redes semánticas. De modo que si nos centramos en la comparación entre un *Topic Map* y un tesoro vemos que este no es sino un caso particularmente simple de *Topic Maps*, donde tan solo existen tres *association types*: (jerarquía, equivalencia y asociación)²¹.

Ciertamente, al partir ambas representaciones de modelos diferentes, el engarce entre ellas no siempre es obvio. Así, no aparece evidente que la diferencia entre los descriptores y los no-descriptores esté más ligada a la pareja [*subject-topic, base name*] o a la de [*topic base name-alternative names*]. Sin duda, la mayor desigualdad respecto a los tesauros viene del momento en que se idearon cada una de las propuestas. Desde su aparición, los tesauros fueron productos centralizados, incorporados a una única organización, y cuya presentación en formato electrónico es solo una ventaja añadida, pero no es requisito imprescindible. Mientras que los *Topic Maps* son productos descentralizados, pensados para mejorar la cooperación entre distintas organizaciones y cuya naturaleza exige tener formato electrónico.

Muestra también interés la relación con la gestión del conocimiento (*Knowledge Management*). Podemos ver entre información y *Knowledge Management* la misma diferencia que existe entre tener un libro y saber la materia de la que ese libro trata. De modo que *Knowledge Management* comprende la generación, codificación y transmisión de la información, donde los *Topic Maps* se establecen como una herramienta de gran valor para esta codificación. Usando reglas que aseguren la inferencia, se puede proceder al *Knowledge Management* basándose en *Topic Maps*.

Los *Topic Maps* muestran asimismo una clara conexión con RDF (*Resource Description Framework*), pues ambos actúan de manera complementaria. La diferencia estriba en que los *Topic Maps* son independientes de que existan o no recursos informativos, alcanzando a permitir la navegación entre conceptos. Mientras que RDF se centra más en la descripción de recursos informativos y no trata la semántica de las relaciones entre metadatos.

Si comparamos los conceptos de los Mapas conceptuales con los que surgen de la normativa *Topic Maps* y, finalmente, ambos con aquellos propios de un tesoro con los que puede establecer alguna correlación o equivalencia, entre las características de ambos sistemas se determina la siguiente tabla comparativa²²:

Mapa conceptual No existe norma	Topic Maps ISO 13250:2000	Tesauros ISO 2788:1986 (o equivalente UNE 50-106-90)
Control del vocabulario. El mapa conceptual no realiza control de vocabulario a priori. En el proceso de aprendizaje <i>profesor</i> y <i>alumno</i> marcan los conceptos y los nodos.	El vocabulario de un <i>topic map</i> viene dado por el documento los documentos que sirven de base para generarlo, y son los términos que van marcando de manera dinámica cuáles son los conceptos subyacentes a cada <i>topic</i>	Los descriptores, y no-descriptores, se marcan a priori.
Dominios. El mapa conceptual se orienta a dominios, pero la validez de cada nodo es algo subjetivo que establece la persona que elabora el mapa conceptual.	<i>Theme/Scope</i> , marca el límite de la validez de las asignaciones de cada <i>topic</i> , límite que va variando con los sucesivos <i>themes</i> en los que van apareciendo los <i>topics</i> , que también pueden ser infinitos.	Relación de Jerarquía (Generalización y enumeración). Organización por campos y organización por indicadores clasificatorios (UNE 50-106-90: 9.3.2.1 y 9.3.2.2)
Tipos de nodos. No existe una tipología de nodos en el mapa conceptual.	<i>Topic types</i> , que representan una relación del tipo clase-instancia, van variando en función del tipo de información (<i>Theme/Scope</i>) que estemos tratando.	Descriptores y no-descriptores de un dominio.
Relaciones. Son arcos etiquetados (verbo o preposición, verbo o nombre, verbo o conector lógico) entre conceptos, nodos.	Las relaciones son etiquetadas sólo por verbos. Son estructuras del tipo [<i>topic</i>] - [verbo] - [<i>topic</i>]. El documento o documentos que sirven de base para generar un <i>topic map</i> son los que van marcando de manera interactiva cuáles son los conceptos subyacentes a cada <i>topic</i> . Además, entre	No existe un conector entre los términos relacionados en un tesauro.
Relaciones. Los mapas conceptuales suelen establecer relaciones del tipo: ' <i>being</i> ', ' <i>having</i> ', ' <i>using</i> ', ' <i>causing</i> ', ' <i>showing</i> ', ' <i>including</i> ', ' <i>similarity</i> ', pero no existe un acuerdo o norma que los fije.	<i>Association types</i> . Los <i>Topic Maps</i> pueden tener un número de relaciones potencialmente infinito.	Relación de Jerarquía Relación de Equivalencia Relación de Asociación
El conjunto 'nodo-enlace-nodo' no puede ser considerado un nodo.	Una <i>association type</i> o, incluso, un <i>topic type</i> pueden ser <i>topic</i> puede ser una <i>association</i> .	Las relaciones que se establecen en un tesauro no suelen ser consideradas descriptores.
Equivalentes ortográficos. No existe esta concreción en el mapa conceptual (trabaja con conceptos con independencia de los términos en los que se exprese).	<i>Topic names</i> , que son las variaciones ortográficas de un mismo término, apodos, acrónimos, ...	Relación de Equivalencia, sólo a través de este tipo de relación es posible contemplar algunos casos similares.
Los nodos no dependen de los documentos en los que puedan aparecer.	<i>Occurrences</i> , casos relevantes del <i>topic</i> que están fuera del propio <i>topic map</i> que se está tratando.	<i>Scope Notes</i> Origen del término Definición del término
El mapa conceptual establece conceptos relevantes en el dominio, pero no términos.	<i>Public subject</i> , que pueden ser tanto los <i>topics</i> que representan al mismo <i>subject</i> , como las variaciones idiomáticas de un <i>topic</i> , quedando todos al mismo nivel.	No descriptor. Nivel diferente. Tesauro multilingüe.
Dominios semántico - conceptuales sólo válidos en los límites de cada mapa conceptual	<i>Facets</i> , expresadas mediante pares atributo-valor, y que pertenecen a la descripción formal del <i>topic</i> y no a la semántico-conceptual.	Facetas (predefinidas en el tesauro)

4 CONCLUSIONES

Entre los límites que han ralentizado las posibles aplicaciones del nuevo paradigma de los *Topic Maps* debemos considerar:

- Su novedad, ya que dan un paso más con respecto a los más conocidos mapas conceptuales de navegación.
- Los diferentes *topic names* que tiene un *topic* ofrecen gran igualdad si los comparamos con las relaciones de equivalencia de un tesoro.
- En la práctica, los tesauros suelen estar contruidos manual o semi-manualmente. Sin embargo, en los *topic maps*, y debido al número de relaciones que pueden tener, el tratamiento manual es imposible, puesto que es potencialmente infinito. Los *topic maps* se relacionan con la web semántica: permiten que el ordenador viaje de un concepto a otro mediante las relaciones semánticas. Por lo cual, se precisan sistemas que permitan la extracción automática de *topic maps*. Por un lado, se deben seleccionar los *topics* mediante reglas de generación, recursos terminológicos y *topic maps* preexistentes. Por otro lado, se deberán crear asociaciones, algo mucho más complejo que necesita algori Topic Mapos de inteligencia artificial²³. Además, el mantenimiento del *topic maps* necesitará de reglas de consistencia, ya que manualmente sólo se pueden tomar muestras, dado su tamaño.
- Una de las razones que explican el éxito de los tesauros es su simplicidad conceptual, con una estructura básica sencilla y limitados tipos de relación. Mientras que en los *topic maps*, el propio motivo que los originó, en su deseo de facilitar la fusión de diferentes recursos, impuso una estructura flexible, con gran riqueza de relaciones. Esta flexibilidad se ha traducido en una cierta ambigüedad terminológica, que junto con las numerosas relaciones lo convierten en un estándar de difícil asimilación inicial. En un tesoro, Paris es un término unívoco, pero en los *Topic Maps*, hechos para fusionar redes semánticas, puede aparecer Paris como ciudad y Paris como héroe mitológico. En este sentido se están haciendo progresos para predefinir, de alguna manera, los *Topic Maps* mediante los *Topic Maps Templates*²⁴.
- La necesidad de mejorar la capacidad de integrar las propiedades de las *association types* dentro del esquema de definición de *Topic Maps*.
- Las restricciones necesarias relativas a su validación y corrección:

- Un *Topic Map* puede presentar miles de *topics* y *associations*, por lo que no resulta práctico revisarlas manualmente para identificar posibles inconsistencias y errores. Ni el estándar ni las DTD relacionadas sugieren ningún mecanismo para solventar estos problemas.
- Las reglas que se han propuesto para asegurar la consistencia se centran principalmente en ver qué *association role types* son validas para determinada *association type*, así: con los *role types* podemos decir que en León <provincia> hay Chopos <árboles>, pero resulta absurdo decir que en un <árbol> hay una <provincia>. Análogamente, se puede decir que las flores <parte de la planta> están en el pino <planta>, pero no puedo decir que el pino <planta> está en las flores <parte de la planta>. Ascendiendo por la jerarquía, mediante *role types* y la propiedad transitiva, podría tener una regla que fuera que el <contenido> está en <continente>, pero no al contrario. Reglas similares se pueden crear para *occurrences*, *scopes* y *topic names*.
- La validación presenta problemas cuando, tras una fusión o una generación de un *Topic Map* a partir del PLN, se puede llegar a *associations* redundantes y hasta incompatibles. Entendemos que *associations* redundantes son aquellas que no aportan información, al poderse obtener los mismos datos mediante otras *associations*. Mientras que incompatibles son aquellas *associations* que mediante la propiedad transitiva pueden indicar diferentes tipos de relaciones referidas al mismo concepto. Si se quieren validar estas *associations* hay que disponer de reglas que comprueben la coherencia del resultado.

Pese a lo cual, los *Topic Maps* presentan evidentes **ventajas**:

- Una mayor riqueza semántica, tanto desde el punto de vista de relaciones entre *topics* como de caracterización del vocabulario.
- Desde una perspectiva muy restrictiva, un *topic map* es un tesoro en cuanto que se va conformando por dominios (*scope* y *themes*), esto es, sirve para un ámbito concreto, con la única diferencia aparente de su carácter dinámico frente a la estaticidad del tesoro. Sin embargo, ofrece la posibilidad de fusionar distintos ámbitos sin que eso suponga la pérdida de información, es decir, el resultado que ofrece no es sólo la mezcla sino un nuevo marco conceptual que engloba a los dominios fusionados.
- El *topic map* establece una serie de relaciones que van marcadas por los verbos que unen a los *topic*, lo que en un tesoro no era posible porque se definían las relaciones

por proximidad (equivalencia, asociación y jerarquía) semántica y conceptual entre sustantivos. Es decir, no necesariamente tienen que pertenecer dos términos a una misma familia semántica o conceptual para que puedan ser relacionados. Lo cual deja abierto el posible número de relaciones.

- Capacidad de organizar recursos informativos de distinto tipo, tanto a escala semántica como formal, algo de indiscutible valor en la web.
- Es una aproximación desde la que se integran fácilmente con glosarios, tesauros, redes semánticas, etc.
- Cierta capacidad de automatización, al estar más próxima al lenguaje natural en las *associative types*.
- Los TOPIC MAPS, pueden dar semántica a elementos que están en el web al organizarlos y describirlos, pero sin modificarlos.
- Perfiles de usuarios: mediante el *scope* y el *theme* permite adaptarse a distintas comunidades compartiendo recursos informativos.
- Navegabilidad e inferencia mediante estructuras semánticas. Lo cual mejora no solo la recuperación de información, sino también la gestión del conocimiento y el mantenimiento de los TOPIC MAPS. En este punto es también interesante recordar la independencia de los recursos informativos que tiene TOPIC MAPS, pudiendo navegar por nodos que no tengan ningún recurso asociado
- Fusión con otras estructuras de conocimiento, permitiendo una gestión descentralizada.

Finalmente, se indican algunos programas informáticos para construir mapas conceptuales:

- *Activity Map* de Time/sSystem Int.: <http://www.timesystem.com/timesystem/software/AMW/>
- *Axon Idea Processor 5.0* de Chan Bok: <http://web.singnet.com.sg/~axon2000/article.h> Topic Maps
- *CMap 2.0 for Macintosh*, accesible mediante gopher://oldal.mannlib.cornell.edu:9570/40/misc/CMap_2.0.hqx
- *CoCo Systems* maker of VisiMap and InfoMap (Lite): <http://www.coco.co.uk>
- *Decision Explorer* (anteriormente *Graphics COPE*) de Banxia Software: <http://www.scotnet.co.uk:80/banxia/demain.h> Topic Maps
- EGLE Magic, de Mind Mapper, que ofrece información por e_mail: MMInfo@emagic.marc.cri.nz
- *Inspiration Software, Inc.* de Inspiration: <http://www.inspiration.com>
- *MindMan* de Michael Jetter: <http://mindman.com>

- *SemNet Research Group* producido por SemNet:
<http://apple.sdsu.edu/logan/SemNet.h> Topic Mapsl
- *SMART Ideas* de SMART Technologies:
<http://www.smarttech.com/smartideas.h> Topic Maps
- *TextVision / TekstNet* hecho por Piet Kommers: <http://130.89.41.31/textvisi.h>
[Topic Maps](#)

Enlaces a contenidos sobre TOPIC MAPS

- <http://www.topicmap.com/> Página muy completa sobre cualquier tema relacionado con TOPIC MAPS
- <http://k42.empolis.co.uk/> Desarrolla software relacionado con TOPIC MAPS. Tiene artículos interesantes
- Topic Maps links. Enlaces a TOPIC MAPS
<http://www.garshol.priv.no/download/> Topic Mapslinks
- www.topicmaps.org página oficial sobre TOPIC MAPS
- www.topicmaps.net links interesantes relacionados con TOPIC MAPS
- http://www.ontopia.net/topicmaps/learn_more
- Links de Ontopia la página también es interesante por el software para TOPIC MAPS que desarrolla
- http://www.sci.sdsu.edu/CRMSE/kfisher_knowrep
- Documentación sobre representación de conocimiento

NOTAS

¹ Los profesores Moreiro González, Marzal García-Quismondo, Beltrán Orenes, y Sánchez Cuadrado, Sonia pertenecen al Departamento de Biblioteconomía y Documentación de la Universidad Carlos III de Madrid. Los profesores Llorens Morillo, y Morato Lara al Departamento de Informática de esa misma Universidad. Todos ellos conforman un grupo de investigación que lleva cinco años investigando la automatización de los tesauros. En la actualidad trabajan en el proyecto REID (Desarrollo de un Tesauro de verbos para entornos de información dinámica. Aplicación del estándar ISO/ICE: 13250: 1999) subvencionado por el Ministerio de Ciencia y Tecnología español en su Plan de I+D.

² Frakes, W. B.; Baeza-Yates, R.- **Information Retrieval. Data Structures and Algorithms**. Upper Saddle River: Prentice Hall, 1992.

³ Peirce, Ch. S.- **Escritos lógicos**. Introducción, selección y traducción de Pilar Castrillo. Madrid: Alianza Editorial, 1988.

⁴ Sowa, J.- **Knowledge Representation: Logical, Philosophical and Computational Foundations**. Pacific Grove: Brooks-Cole, 2000.

⁵ Walsh, N. y Muellner, L.- **DocBook: The Definitive Guide**. Sebastopol: O'Reilly, 1999.

⁶ De acuerdo con: http://users.edte.utwente.nl/lanzing/cm_home (visitado 04.04.01).

⁷ Anderson, L., Ditson, L., Ditson, M.- **Computer Based Concept Mapping: Promoting Meaningful Learning in Science for Students with Disabilities** [documento www]. Disponible desde 1998 en: <http://www2.edu.org/NCIP/library/ot/zeitz>

⁸ El término *Topic Navigation Maps* fue acuñado por Michel Biezunski quien diseñó y editó originalmente la arquitectura, y fue aceptado en 1996 por el grupo de trabajo SGML dentro de la ISIO. El texto completo del estándar ISO/ICE 13250:1999 sobre *Topic Maps* se puede encontrar en <http://www.infoloom.com/Topic Mapsstands> [consultado el 28 de febrero de 2002].

⁹ Pepper, S.- Euler, Topic Maps, and Revolution, <http://www.infoloom.com/Topic Mapssample/pep4> [consultado el 16 de julio de 2002]. Su éxito se tradujo en la aprobación del estándar *Topic Maps*: ISO 13250. International Organization for Standardization, ISO/IEC 13250.- **Information Technology-SGML Applications- Topic Maps**. Geneve: ISO, 2000.

¹⁰ Rath, H. H.- Technical Issues on Topic Maps, en **MetaStructures 99**.

¹¹ "V" topic map browser, J. Algermissen: <http://www.topicmapping.com/v>; Ontopia Omnigator, de Ontopia, tiene [demo](#) y shareware; K42, de Empolis, tiene una [demo](#) y una muestra shareware en Java.

¹² Moreira, J.A.; Sánchez Cuadrado, S.; Morato, J.- **Panorámica y tendencias en topic maps** [en línea], en Rovira, C., Codina, L. (dir.).- **Documentación digital**. Barcelona: Sección Científica de Ciencias de la Documentación. Departamento de Ciencias Políticas y Sociales. Universidad Pompeu Fabra, 2002. <<http://www.documentaciondigital.org>> [Consulta: 11 abr. 2003]. ISBN 84-88042-39-6.

¹³ Según Pepper, el concepto de *subject* fue definido en primera instancia como un constructo mental que surge de un objeto: "el corazón invisible de cada *topic* es el *subject* que su autor tenía en mente cuando lo creó". Pepper, S.- **"Euler, Topic Maps, and Revolution"**, <http://www.infoloom.com/Topic Mapssample/> [consultado el 28 de febrero de 2002].

¹⁴ Los signos de un mapa conceptual se definen en Arquitectura SGML, por lo que esta norma internacional toma la forma de un *Architecture definition document* expresado de conformidad con el Anexo Normativo A.3 de ISO/IEC 10744:1997, el *SGML Architectural Form Definition Requirements* (AFDR). La definición formal de la notación de un mapa conceptual se expresa como un meta-DTD.

¹⁵ Definida por la arquitectura HyTime en ISO/IEC 10744:1997.

¹⁶ Tal como se explica en el Anexo K de SGML (1997), también conocido como WebSGML.

¹⁷ Rath, H. y Pepper, S.- Topic Maps at work, en **The XML Handbook** - 2nd Edition. New Jersey: Prentice Hall, 1999. <http://www.topicmaps.com/content/resources> [consultado el 9 de julio de 2002].

¹⁸ Dos conceptos pueden conectarse mediante una asociación, y pueden también conectarse en virtud de participar en un acontecimiento común. Además, los objetos informativos pueden tener propiedades, así como valores de esas propiedades, asignados a ellos externamente. Estas propiedades son denominadas "*facet types*": La palabra faceta se utiliza aquí de forma metafórica: una faceta es una propiedad de un grupo de documentos que puede ser usada para crear una perspectiva suya.

¹⁹ Novak, J.D.- Clarify with concept maps: A tool for students and teachers alike, en **The Science Teacher**, 1991, 58, n° 7: 45-49.

²⁰ Le Grand, B., y Soto, M.- Visualisation of the semantic web: Topic Maps Visualisation, en **Information Visualisation 6th International Conference. 10-12 July 2002**: 203-225.

²¹ Moreiro, J. A.; Sánchez Cuadrado, S.; Morato, J.- Panorámica y Tendencias sobre *Topic Maps* (en línea), en Rovira, C. y Codina, L. (dir.).- **Documentación digital**. Barcelona: Sección Científica de Ciencias de la Documentación. Departamento de Ciencias Políticas y Sociales. Universidad Pompeu Fabra, 2002. ISBN 84-88042-39-6: <http://www.editaweb.com/docdigitalinfo/>

²² Existen algunos conceptos propios de la norma *Topic maps* que no guardan ninguna correspondencia con los conceptos de los tesauros: como es el caso de *association*, *association role*, que tiene una larga trayectoria en procesamiento del lenguaje natural pero no en los sistemas de los tesauros, o el de *occurrence role* y que, por lo tanto, no caben en el esquema comparativo que aquí mostramos. Moreiro, J. A.; Llorens, J.; Marzal, M. A.; Morato, J.; Sánchez, S.; Beltrán, P.- Construcción automática de *Topic Maps*, en **Actas del XX Congreso Brasileiro de Biblioteconomia, Documentação e Ciencia da Informação**. Dimensão Humana, Política e Econômica da Informação. Fortaleza-Brasil, 2002 [/Acesso/Construcción.h Topic Maps].

²³ Rath, H. H.. Technical Issues on Topic Maps. **Proceedings of Metastructures 99 Conference**. Alexandria: GCA, 1999.

²⁴ Rath, H.- Topic Maps: Templates, topology and type hierarchies, en Sperberg-McQueen, C. M.; Usdin, B. T. (eds).- **Markup Languages**. Cambridge (Ma): MIT Press, 2000.

ABSTRACT

The noun growth of the number of thesauri has not served to respond to the necessity to work in multidisciplinary surroundings. In order to respond to this situation, it was fomented, at a first moment, the fusion of thesauri with object to adapt the preexisting ones to the necessities raised by the new dominions. Later, in the eagerness to take care of a world of changing information indeed and in growth, one began to work in the conceptual navigation maps. Until ending at the Topic maps arisen from the necessity to fuse indexes to incorporate therefore the utility of the hyperconnections. Its success was shaped in the norm ISO/ICE 13250:1999, that motivated the hypothesis of its use in the thesaurus elaboration, process in which the contradictions between both systems were pronounced. Reason why the essential conceptual elements in the architecture of Topic Maps are analyzed to find their costory in thesauri. This comparative study allows to indicate like limits of Topics maps, their conceptual indefinición, its innumerable relations and their ambiguity, which does not hide its advantages like a greater semantic wealth, the attainment of a new conceptual frame for fused dominions, the associations determined by verbs, and the capacity to organize informative resources of different type.

KEYWORDS: Topic-Maps. Thesauri. Conceptual Maps. Representation of the information. Automatic Processing of the Information.

Originals recebidos em 14/06/2004