

Los Topic Maps y su relación con las redes sociales

Lic. Yeter Caraballo Pérez¹

RESUMEN

Se realizan algunas consideraciones teóricas sobre el auge de las especificaciones como XML Topic Maps. Se exponen las principales definiciones y características de los Topic Maps. Se identifican los aspectos que relacionan a los Topic Maps con las redes sociales, a partir de los avances que estos ofrecen en materia de gestión del conocimiento y con respecto a la consolidación de la Web semántica.

Palabras clave: Topic Maps, redes sociales, gestión del conocimiento, World Wide Web.

ABSTRACT

Some theoretical considerations on the increase of specifications such as XML Topic Maps are made. The main definitions and characteristics of Topic Maps are exposed. The aspects that relate Topic Maps to the social networks are identified through the advances they offer in connection with management knowledge and the semantic Web consolidation.

Key words: Topic Maps, social networks, management knowledge, World Wide Web.

Copyright: © ECIMED. Contribución de acceso abierto, distribuida bajo los términos de la Licencia Creative Commons Reconocimiento-No Comercial-Compartir Igual 2.0, que permite consultar, reproducir, distribuir, comunicar públicamente y utilizar los resultados del trabajo en la práctica, así como todos sus derivados, sin propósitos comerciales y con licencia idéntica, siempre que se cite adecuadamente el autor o los autores y su fuente original.

Cita (Vancouver): Caraballo Pérez Y. Los Topic Maps y su relación con las redes sociales. Acimed 2007;16(4). Disponible en:

http://bvs.sld.cu/revistas/aci/vol16_4_07/aci091007.htm [Consultado: día/mes/año].

El desarrollo inmediato de la Web y el incremento cada vez más significativo del volumen de la información digital han propiciado que se creen especificaciones como *XML Topic Maps (XTM)*, un modelo que provee una gramática para representar la estructura de los recursos de información utilizados para definir temas (topics), y las relaciones entre los temas. Es decir, nombres, recursos y las relaciones entre ellos se establecen mediante una serie de características de contenidos abstractos, que denominadas "topics". A su vez, los *topics* se caracterizan por una serie de elementos como los "scopes", que no son más que los ámbitos, campos o contextos limitados en los cuales se consideran. Cuando uno o más documentos interrelacionados emplean esta gramática, constituyen lo que se denomina, según la literatura estudiada, un mapa temático o topic map.

Lamarca Lapuente (2007), en su tesis doctoral "*Hipertexto: El nuevo concepto de documento en la cultura de la imagen. XML Topic Maps*" afirma que un *Topic Map* (TM) tiene como finalidad normalizar los elementos y la notación utilizada para estructurar la información mediante la construcción de una red de enlaces semánticos que relacionen diferentes recursos informativos.¹⁻³

La especificación XML Topic Maps (XTM) describe un modelo abstracto y una gramática para el intercambio de Topic Maps basados en el Web. Esta especificación es una creación de TopicMaps.Org Authoring Group.

TOPIC MAPS

Los TM son un estándar que se introduce progresivamente con un gran éxito en la World Wide Web y que posibilita la navegación entre conceptos y la inferencia, y esto constittación gráfica que permite el acceso y navegación por los recursos de un campo del conocimiento mediante los llamados mapas temáticos, es decir, un mapa conceptual aplicado a la World Wide Web.

En esta línea, *TopicMaps.org* (<http://www.topicmaps.org/>), un consorcio independiente que desarrolla la aplicación del modelo de mapas temáticos especificado en la norma ISO 13250:2000 para la World Wide Web, establece una serie de especificaciones para XML. Todas las versiones de las distintas especificaciones XTM son públicas.

Los elementos principales de los TM son: los temas (topics), las ocurrencias (occurrences) y las asociaciones (associations), estas últimas representan sólo una parte del *topic*. Para deshacer la ambigüedad que puede crear la simple aparición de un nombre como "Pedro" (topic), este topic puede asociarse con otros datos o documentos, por esto un *topic* es el elemento principal de un TM y es el término que expresa determinado concepto o idea (subject).²

Las características principales de un TM son:

- *Su denominación (<name>)* Aunque un topic puede presentar varias denominaciones, el nombre debe estar representado obligatoriamente por una forma base (<baseName>). El "base name" representa la forma normal de referirse al topic. Es complementario del identificador interno (id). Los "base name" deben de ser únicos en un dominio determinado bajo un ámbito (scope) determinado. Cuando se fusionan dos TM, se tratan los distintos "base names" de los "topics" que hacen mención a un mismo "subject" (concepto). Y para ello, se normalizan los "base names" mediante los "public subject", un recurso que identifica al "subject" de forma no ambigua. Pero, 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 solicite una lista.
- *Sus apariciones (<occurrence>)*. Son enlaces a recursos de información, como por ejemplo, una página Web. Las "occurrences" son elementos opcionales que son relevantes a un "topic" determinado. Un topic puede tener cero o muchas "occurrences", y éstas pueden ser de muchos tipos: una cita de texto, una definición, un artículo, etcétera. Cada uno de los diferentes tipos de documento puede agruparse mediante "occurrence roles", por ejemplo, página Web, imagen,

diccionario, entre otros. La mayoría de estas ocurrencias suelen ser externas al TM, como ocurriría entre los índices documentales y los tesauros. A cada “occurrence”, le corresponde un único enlace. Existen dos tipos de “occurrences”: un dato expreso (resourceData) y un enlace a un recurso externo (resourceRef).

- *Sus roles de asociación (<association>)*. La "association" es un enlace que establece una relación entre dos o más topics. En las siguientes frases, pueden encontrarse los siguientes topics: álamo, chopo, madera, árboles, leña, madera, chopos, ribera, bosque y álamos, los términos de unión entre los topics forman la asociación.

Una de las características de los TMs que permiten la posibilidad de adaptarse a diferentes perfiles de usuarios. Esto se realiza por medio de dos tipos de filtrado: el "scope", que filtra diferentes topics y "themes" (temas) para una determinada comunidad, y las facetas (facets), que se ocupan de filtrar recursos informativos, por ejemplo, ofrecer a determinado usuario sólo los recursos informativos que estén en un idioma determinado.

Para la creación de TM existen muchos y variados métodos que pueden simplificar su creación. Se puede partir del procesamiento del lenguaje natural mediante frases en lenguaje libre para luego identificar estructuras verbales que pueden asociarse de forma automática con determinados “association types”. Sin embargo, este método presenta muchos problemas debido a la polisemia, riqueza del lenguaje y ambigüedad. Más práctico resulta, mediante las propiedades de las asociaciones, obtener nuevas asociaciones distintas a las originales, o utilizar la fusión de recursos.

Los TM representan una de las principales propuestas para la presentación de la Web semántica y la puesta en práctica de mapas temáticos y los mapas conceptuales de la Web. Sin embargo, existen todavía muchos problemas por resolver, porque no hay que olvidar que un TM puede tener un gran número de asociaciones de diferentes tipos (association types, roles, occurrences, etcétera). Para mostrar los TM en la Web, se han propuesto y se utilizan varios tipos de representaciones, entre las que destacan las tablas (basadas en tablas con enlaces de referencias cruzadas), los árboles (modelos jerárquicos de datos con hiperenlaces), los gráficos (como RDF, TMs) y las corrientes de texto desestructurado. A la hora de analizar estas representaciones, se debe considerar que muestren tanto información local de los topics que interesan al usuario, como información sobre la localización de esos topics en el conjunto del TM.⁴

La perspectiva de los TM, desde las redes sociales, es que estos se encuentran en la práctica implementados en el Web, y propician:²

La adaptación a distintos perfiles de usuarios. Un TM ofrece la posibilidad de adaptarse a diferentes perfiles de usuarios. El modo de realizarlo es considerando dos tipos de filtrados. Por un lado, con el “scope” que filtra diferentes topics y themes (temas) para una determinada comunidad. Por otro, las facets que se encargan de filtrar recursos de información, por ejemplo, ofrecer a determinado usuario sólo los recursos informativos que estén en español.

Scopes y Themes: Los “scopes” y los “themes” representan el ámbito en el que las afirmaciones realizadas sobre algún “name”, “occurrence” o “association” de un “topic”

tienen coherencia y validez. Un “theme” es la forma en que se denomina un determinado “scope”.

Facets: Las “facets” proporcionan un mecanismo para asignar pares de propiedad-valor a los recursos de información, por ejemplo, considerar sólo documentos en un idioma determinado con una aplicación determinada (experimental, práctica...).

La relación con otros recursos para representar el conocimiento (KR). Los TM se proponen para modelar redes semánticas. Por ejemplo, si nos centramos en la relación entre un TM y un tesoro observamos que los tesauros no son más que un caso particularmente simple de TM, donde tan solo existen tres “association types” (jerarquía, sinonimia y relacionado). Así, la relación de jerarquía es una “association type” que podría denominarse “es un” o “es una clase de”.

Ciertamente, al diferir el modelado inicial de TM y tesauros, el engarce no es siempre obvio. Por ejemplo, no parece tan evidente si la diferencia entre descriptores y no-descriptores está más ligada a la pareja [*subject-topic, base name*] o a [*topic base name-alternative names*]. Otras diferencias con los tesauros están más unidas al momento histórico en que cada propuesta tuvo lugar; así los tesauros son normalmente productos muy incorporados a una única organización (centralizados) y cuya presentación en formato electrónico es sólo una característica extra pero no un requisito. Los TM son productos descentralizados que mejoran con la cooperación de distintas organizaciones y que tienen esencialmente un formato electrónico.

Otro aspecto interesante en relación con la gestión del conocimiento (KM o Knowledge Management). Si se analiza la diferencia entre información y KM es la misma que entre tener un libro y saber de la materia que trata el libro. Esto es, KM comprende la generación, codificación y transmisión de la información; TM una herramienta muy valiosa para esta codificación. Mediante las reglas que aseguran la inferencia se puede llegar a realizar un KM basado en TM.

En resumen, la utilización de los TM ofrece una serie de avances en el camino hacia la Web semántica y la gestión del conocimiento y, por tanto, en la consolidación de las redes sociales. Estos son:

- Permiten dotar de semántica a muchos elementos y recursos del Web, para poder describirlos y organizarlos.
- Permiten distintos perfiles de usuarios mediante la utilización de las características “scope” y “theme” para adaptarse a las distintas comunidades que pueden compartir los recursos informativos
- Permiten la navegación y la inferencia mediante estructuras semánticas.
- Permiten una mejor gestión del conocimiento y mejoran los sistemas de búsqueda y recuperación de información.
- Pueden combinarse con otras estructuras de conocimiento para facilitar una gestión descentralizada.
- Se adaptan a un gran número de recursos de información distintos.

REDES SOCIALES

Uno de los componentes fundamentales de la Web semántica son las firmas digitales. Estas se definen como “bloques de datos cifrados que las computadoras y agentes de software podrán utilizar para verificar que la información adjunta fue proporcionada por una fuente fiable”. Al aplicar un algoritmo de cifrado sobre un documento o mensaje a firmar, se obtiene un bloque de datos que representa un resumen del documento o mensaje. La firma digital se obtiene mediante un *Sistema de criptografía de clave pública*, que se basa en el empleo de dos tipos de claves: pública y privada, todo lo que se cifre mediante una clave pública sólo puede descifrarse por medio de su correspondiente clave privada y viceversa.

Las ventajas que confiere la firma digital a los recursos, documentos o mensajes del Web son principalmente:

- *Identificación*. Posibilidad de determinar la identidad del emisor o autor del recurso.
- *Integridad*. Facilidad para detectar la manipulación o alteración ilícita del recurso.

La privacidad de los datos personales que se manejan en Internet es una preocupación constante para empresas, gobiernos, medios de comunicación y el público en general. Para solucionar este problema, surgieron grupos de trabajo e iniciativas dirigidas a estandarizar estos problemas como los grupos de trabajo XML Encryption y XML Signature, este último fruto del esfuerzo conjunto entre el W3C y el IETF (Internet Engineering Task Force), quienes han desarrollado una especificación, con sintaxis XML, para la representación de firmas digitales en recursos Web que alcanzó el grado de recomendación del W3C en febrero de 2002.

Otro esfuerzo orientado a la seguridad es P3P, o *Plataforma de preferencias de privacidad (Platform for Privacy Preferences)* esta posibilita:

- Desarrollar herramientas y servicios que ofrezcan a los usuarios un mayor control sobre la información personal que se maneja en Internet y, al mismo tiempo, aumentar la confianza entre los servicios Web y los usuarios.
- Mejorar el control del usuario al colocar políticas de privacidad donde los usuarios pueden encontrarlas, en un formato en el que los usuarios pueden entender y, lo más importante, con la posibilidad de que el usuario actúe sobre lo que ve.
- A los usuarios Web, facilidad y regularidad a la hora de decidir si quieren o no, y bajo qué circunstancia, revelar información personal.

La solución puede estar en el concepto de una *Web de Confianza*, que basa su funcionamiento en la confianza o fiabilidad entre usuarios y agentes de software en la red, definido como “el grado en que un agente considera un aserto como verdadero para un contexto determinado”. Cada usuario, identificado por un URI (*Universal Resource Identifier*), expresaría su grado de confianza o desconfianza sobre otros individuos, que a su vez harían lo mismo sobre otros, lo que daría como resultado extensas e interoperables redes de confianza procesables por agentes inteligentes.

Para la Web semántica se han puesto en práctica aplicaciones que propician el análisis de redes sociales que toman como premisa un vocabulario RDF para describir

información sobre los individuos y sus relaciones en una red social, conocido como FOAF (Friend *OF A Friend*), que se extiende con el objetivo de expresar el nivel de confianza entre los individuos de la red para finalmente mostrar su implementación práctica en áreas como el correo electrónico y la mensajería instantánea.

Una red es una forma abstracta de mostrar una serie de sistemas y, en general, casi todos los sistemas complejos. Las redes, al fin y al cabo, están compuestas de nudos, que se llaman habitualmente nodos, y de enlaces entre ellos, que se llaman aristas, si es que son flechas que van de un nodo al otro con un sentido definido; o bien arcos, si es que la relación es recíproca, o por decirlo de otro modo, las flechas tienen puntas en los dos extremos.⁵⁻⁷

También es una red un grupo de páginas Web que utilicen hiperenlaces para referirse unas a otras. Si esas páginas Web están escritas por una sola persona, o son directamente páginas Web personales, los enlaces pueden reflejar una relación social entre los creadores de la Web, que se podría expresar vagamente como *conoce-a*. El escritor de una Web que ha incluido un enlace a la segunda Web, *conoce-al* autor de esa segunda Web. Las redes sociales son también redes complejas, aunque emplean una terminología ligeramente diferente: los nodos son agentes, porque hacen algo, mientras que las aristas o arcos expresan, habitualmente, una relación social como *conoce-a*, *es-amigo-de*, o *han-comido-spaghettis-juntos*.

Las redes sociales están cobrando hoy día mucha importancia por los siguientes motivos:

- Todos pertenecemos a una o más de ellas.
- Las redes sociales son los vehículos de la influencia y el poder en las organizaciones.
- Las propiedades inherentes a las redes determinan sus posibilidades. Existen acciones que se pueden hacer en una red y otras que no.
- El advenimiento de Internet, la red por excelencia, ha propiciado la aparición de miles de redes, con muchos miembros cada una.

Por eso, el estudio del análisis y la presentación de las redes de interacción social ha empezado a recibir una cierta atención. Un punto importante de las redes y las sociales son simplemente una más entre las muchas redes posibles; es que existe una teoría matemática, la teoría de grafos, que permite estudiar el comportamiento de las redes y conocer sus propiedades.

El reducir las relaciones sociales a un grafo, es decir, a un conjunto de nodos con unas relaciones explícitas entre ellos, permite hacer una serie de estudios, de los cuales se pueden extraer conclusiones desde simples —cuántas personas intermediarias harían falta para conseguir el número de móvil de una determinada persona— hasta complejas —quién es el agente con más influencia en una red social.

Más aún, como su propio nombre indica, la teoría de grafos tiene un componente esencialmente gráfico, fácilmente representable. Y, lo que es mejor todavía, sus implicaciones son relativamente fáciles de comprender al observar un grafo. Todo esto la hace ideal como metáfora visual para la representación de la información de los

sistemas en la red, motivo por lo cual la mayoría de las representaciones sobre redes sociales utilizan este paradigma.

Construir un grafo de una red social y mostrarlo puede ser un factor de decisión importante. En todas las redes existen nodos que acumulan enlaces, mientras que otros apenas están ligados a los demás. Obviamente los nodos preponderantes, con más enlaces, son también los nodos más influyentes. No siempre es fácil ni, sobre todo, rápido detectarlos.

El análisis de redes sociales es una metodología que busca, mediante la aplicación de modelos extraídos de la *Teoría de redes sociales*, predecir el comportamiento de una red social y aproximar las estrategias de los nodos que la componen. Para eso, es necesario disponer de la información sobre qué nodos existen y cómo están conectados entre ellos. Con esto, es posible construir el grafo, representarlo y estudiar los indicadores que nos permitan valorar la estructura de la red, así como conocer qué nodos tienen un papel decisivo y cuáles no.

La relación entre los TM y las redes sociales es un campo que en la literatura no se ha consolidado del todo y se exponen separados indistintamente; precisamente el establecimiento de dicha relación es lo que le aporta un carácter novedoso a este estudio y al análisis y visualización de las redes sociales en el Web.

Existen varios indicadores que permiten analizar el trazo de la red social de un grupo de bloggers, y sólo por medio de Internet se obtendrían datos de distintas categorías, como los siguientes:

- Los enlaces que los nodos hacen desde sus páginas Web. Como sus páginas son en su mayoría bitácoras, podríamos además clasificar por épocas a quién han enlazado más o menos, además de *via archivos* a quienes han enlazado de forma permanente en cada época.
- Las colaboraciones en medios, libros u otras Web.
- Si publican libros, qué autores han publicado en esas mismas editoriales en distintos años, quienes seleccionaron los contenidos, etcétera.
- Las conferencias en las que cada uno de los miembros ha participado como ponente en cada uno de esos períodos, con quiénes ha coincidido en las mesas y quiénes las organizaban y patrocinaban.
- Los órganos institucionales a los que han pertenecido (desde directivas de asociaciones a órganos de administraciones o empresas) y, por supuesto, con quiénes han coincidido y con qué frecuencia en cada una de ellas.
- Las empresas que participan, con las que colaboran o en las que trabajan y junto a quién en cada momento.
- Los proyectos colectivos en los que han tomado parte —voluntariado, jurados de concursos, etc— y junto a quiénes.

Al mismo tiempo, se podría obtener una lista de las referencias a sus parejas, familiares, compañeros de colegio, promoción e, incluso, vecindario, con vista a establecer relaciones de largo recorrido. Todo esto se debería hacer también, al menos, para cada uno de los contactos que están a uno o dos grados de separación de los nodos estudiados.

REFERENCIAS BIBLIOGRÁFICAS

1. Lamarca Lapuente MJ. Hipertexto, el nuevo concepto de documento en la cultura de la imagen. Tesis doctoral. Universidad Complutense de Madrid. Disponible en: <http://www.hipertexto.info> [Consultado: 7 de mayo de 2007].
2. Moreiro JA, Sánchez Cuadrado S, Morato J. Panorámica y tendencias en Topic Maps.Hipertext.net 2003(1). Disponible en: <http://www.hipertext.net/web/pag229.htm> [Consultado: 7 de mayo de 2007].
3. Silio Martínez T. Consideraciones acerca de la intertextualidad: los Topic Maps como redes semánticas en forma de meta-índices tópicos. Disponible en: <http://eprints.rclis.org/archive/00004100/01/OCSTSilio.pdf> [Consultado: 7 de mayo de 2007].
4. Rodríguez Perojo K, Ronda León R. Web semántica: un nuevo enfoque para la organización, la recuperación de la información el Web. *Acimed* 2005;13(6). Disponible en: http://bvs.sld.cu/revistas/aci/vol13_6_05/aci03605.htm [Consultado: 7 de mayo de 2007].
5. Merelo Guervós JJ. Redes sociales: una introducción. Disponible en: <http://revista-redes.rediris.es/webredes/talleres/redes-sociales.pdf> [Consultado: 10 de mayo de 2007].
6. Dürstelner JC. Visualización de redes sociales.2003. Disponible en: <http://www.infovis.net/printMag.php?num=136&lang=1> [Consultado: 7 de mayo de 2007].
7. Deugarte. Análisis de redes sociales. Disponible en: http://www.deugarte.com/wiki/contextos/An%C3%A1lisis_de_redes_sociales [Consultado: 10 de mayo de 2007].

Lic. Yeter Caraballo Pérez. Centro de Información. Instituto de Ciencia Animal (ICA). Ministerio de Educación Superior (MES). Carretera Central Km. 471/2 Apartado 24 CP 32700. San José de las Lajas, La Habana, Cuba. Correo electrónico: jeter@ica.co.cu

¹Licenciada en Bibliotecología y Ciencia de la Información. Centro de Información. Instituto de Ciencia Animal. Cuba.

Ficha de procesamiento

Términos sugeridos para la indización

Según DeCS¹

ALMACENAMIENTO Y RECUPERACIÓN DE LA INFORMACIÓN, WWW.
INFORMATION STORAGE AND RETRIEVAL; WWW.

Según DeCI²

ALMACENAMIENTO Y RECUPERACIÓN DE LA INFORMACIÓN; WWW.
INFORMATION STORAGE AND RETRIEVAL; WWW.

¹BIREME. Descriptores en Ciencias de la Salud (DeCS). Sao Paulo: BIREME, 2004.

Disponible en: <http://decs.bvs.br/E/homepagee.htm>

²Díaz del Campo S. Propuesta de términos para la indización en Ciencias de la Información. Descriptores en Ciencias de la Información (DeCI). Disponible en:

<http://cis.sld.cu/E/tesauro.pdf>