

# Tendencias en la indización y recuperación temática

CATALINA NAUMIS PEÑA

*Instituto de Investigaciones Bibliotecológicas,  
y de la Información UNAM*

*El epicentro de la cultura se ha trasladado al mundo virtual de Internet, y no tiene sentido protestar por ello, porque la historia no tiene libro de reclamaciones. (Savater, 2012)*  
*En tiempos pasados el gran problema era que la información llegaba a cuentagotas y hoy lo hace en aluviones. (Savater, 2014)*

## INTRODUCCIÓN

**S**in duda, la gran tendencia en la indización y recuperación temática es el estudio de la organización de contenidos en Internet, con la incorporación de la web, cada vez más producida con base a estructuras que tratan de emular el trabajo de organización realizado en las bibliotecas tradicionales y superarlo a través de una extraordinaria fluidez para cambiar las relaciones usuario e información. Además de esa gran fluidez se está logrando la interoperabilidad entre sistemas de información para indizar, buscar y recuperar información. El objetivo de este trabajo es analizar la indización humana y la recuperación de información marcada ahora por la fluidez e interoperabilidad de la información navegando en la web.

Nunca ha sido suficiente con tener acceso o poseer información, es necesario, saber hacer un uso adecuado de esa información para resolver o solucionar problemas o situaciones reales. El cúmulo de

información disponible es de tal magnitud que para ser manejable y transformarla en conocimiento, en tiempo limitado y situaciones muy concretas exige una respuesta más relacionada con el contexto donde se origina la información. Conocimiento es información en acción y desde esta perspectiva, la organización del conocimiento actúa como una capa superior de inteligencia que se superpone a los sistemas tradicionales de organización de la información.

El acceso a distancia disponible en la actualidad permite obtener conocimiento como por ejemplo a través de la asistencia a una operación o experimento desde el otro extremo del mundo donde se realiza. (David y Foray, 2002). El reto es indizar y preparar sistemas para que en el futuro se pueda recuperar ese conocimiento tantas veces como se requiera.

Desde el punto de vista de los desarrolladores de tecnología la solución está en disponer de herramientas potentes y sumamente rápidas, que nos ayuden a tomar decisiones proactivas, y conducidas por un conocimiento acabado de la información. Estas herramientas exploran las grandes bases de datos de forma eficiente e incluso permiten buscar patrones ocultos o información predecible de series temporales.

Sin embargo, como se observa en el ejemplo anterior la organización necesaria para recuperar estas grandes oportunidades de aprendizaje no se realiza sólo con la tecnología que permitió vivir la experiencia y conservarla. Alguien humano tiene que seleccionar, ordenar e indizar e incorporarlo a sistemas de los que se espera ese tipo de respuesta.

La oferta de la tecnología facilita muchísimas cosas, pero algunas de las respuestas acerca de conocimiento científico o tecnológico para situaciones específicas y especializadas se dan a través de los sistemas de información implementados por especialistas del campo de la Bibliotecología y la Información. La pregunta es: ¿La Bibliotecología que ofrece como respuesta a esa inquietud de los desarrolladores de tecnología que aseguran recuperar conocimiento a través del procesamiento automático de la información?

## INFORMACIÓN Y TECNOLOGÍA QUE LA SOPORTA

Sin duda, las computadoras nos ayudan a transformar datos en información, sin embargo aún en el futuro, es improbable que de forma automática, se analicen los documentos en el contexto que se generan. Esta es una buena razón para distinguir entre la información y la tecnología que la soporta.

El texto tanto impreso como audiovisual es el producto del conocimiento que se plasma explícitamente en forma de recursos de información. En general existe consenso en declarar el conocimiento como la combinación de información, personas, tecnología y contexto en la medida que producen acciones. Es decir, el conocimiento se produce con la capacidad de razonamiento, experiencia, interpretación y adaptación del hombre, que no puede ser analizado por una computadora sobre la información plasmada en un documento o en los recursos de información que se suman unos a otros conformando grandes cúmulos de datos o información.

En la web se puede navegar, pero para alimentarse es necesario buscar y pescar peces gordos y para ello hay que parar en algún lugar donde existan garantías de una buena pesca y por supuesto contar con la tecnología para ello. Entre estos lugares de buena pesca o de los que interesa mencionar en este trabajo están aquellos donde se transfiere el conocimiento, como los sistemas bibliotecarios y las bases de datos, que seleccionan, organizan e incluyen, la creación de nuevas ideas o de nuevos procesos o la recopilación de conocimientos relacionados. El acceso a estos centros de conocimiento se produce tanto por Internet como por la web, agregando a partir de la segunda los datos enlazados para realizar las operaciones de hipertextualidad que caracterizan a la web.

La misión del Grupo *Incubadora de Datos Enlazados Bibliotecarios* del W3C (*World Wide Web Consortium*) comisionado entre mayo de 2010 y agosto de 2011, ha consistido en:

Ayudar a aumentar la interoperabilidad global de los datos bibliotecarios en la web, reuniendo a personas implicadas en las actividades de la web semántica - centradas en los datos enlazados- en un sentido

amplio en el campo de las bibliotecas, a partir de iniciativas ya existentes, e identificando vías de colaboración futuras. (Informe del grupo W3C Incubator, 2011)

En el documento mencionado en el párrafo anterior, se identifican un conjunto de recursos útiles para crear o consumir datos enlazados en el campo de las bibliotecas. Un tema central en el documento generado por el grupo es tratar de transmitir al medio bibliotecario como enlazar los datos. Es decir es el puente entre la tecnología a aplicar en los sistemas bibliotecarios y las tradiciones profesionales en el campo.

El informe del Grupo Incubadora de Datos Enlazados Bibliotecarios: *Conjuntos de datos, Vocabularios controlados y Conjuntos de Elementos de Metadatos* fue publicado en octubre de 2011 y la consulta actual al MeSH o Encabezamientos de la base de datos de la Biblioteca Nacional de Medicina de Estados Unidos o al *Agrovoc* muestran cómo están enlazados con otros recursos como por ejemplo el Fichero de nombres Personales de la Biblioteca del Congreso y ofrecen unas posibilidades de recuperación a través de relaciones de cualquier tipo entre información. Ya no son necesarios los algoritmos de antaño para cruzar información en situaciones concretas.

Uno de los aspectos importantes de este trabajo del grupo incubadora es que está libre de derechos y en el marco de la W3C. Entre los datos enlazados se distinguen los tres grupos de recursos que están identificados por este grupo incubadora: los conjuntos de datos, los vocabularios controlados y los conjuntos de elementos de metadatos.

Los conjuntos de datos se refieren a las descripciones de objetos como los libros de una biblioteca o las fotografías de una fototeca y en ellos se distinguen las entidades que conforman cada registro y el valor que tiene cada uno de estos elementos.

En el mundo bibliotecario el equivalente a un conjunto de datos es una colección de registros bibliográficos, en la que cada declaración (usando la terminología de cómputo) consiste en un elemento (*atributo o relación*) de una entidad, y un *valor* para cada elemento. Los elementos que se usan suelen estar seleccionados de un conjunto de elementos estándar, como *Marc21* o *Dublin Core (DC)*. Los valores

de los elementos se toman bien de vocabularios controlados como las Listas de Encabezamientos de Materia, tesauros o son simplemente texto libre.

El documento sobre los datos enlazados ejemplifica de la siguiente manera la relación tecnología y tradiciones bibliotecarias:

- Un registro de un conjunto de datos para un determinado libro tiene un elemento *subject*, tomado de *Dublin Core*, y un valor para esa materia tomado de la LCSH.
- El mismo conjunto de datos puede contener registros de autores como entidades de clase 1 que están enlazadas con sus libros, y descritas mediante elementos como *name*, procedente de FOAF (*Friend of a Friend*, ontología usada para describir personas y sus relaciones con otras personas y con recursos web), o de registros de autoridades bibliográficas.
- Un conjunto de datos puede ser auto-descriptivo al contener información sobre sí mismo como una entidad distinta, por ejemplo al incluir los elementos fecha de modificación y responsable de mantenimiento tomados de *Dublin Core*.

Los vocabularios controlados definen recursos (por ejemplo, temas, estilos artísticos o autores) que se usan como valores. El caso típico de un vocabulario controlado no define recursos bibliográficos como libros o películas o fotos sino más bien conceptos relacionados con ellos (personas, lenguas, países) Estas herramientas de tipo lingüístico son componentes estructurales que sirven para rellenar de contenido los metadatos. Para ser útiles en el enlazado de datos, han de tener URI (*Uniform Resource Identifier*) accesibles a través del protocolo HTTP (*Hipertext Transport Protocol*) asignados para cada elemento descriptivo o de contenido. Las referencias URI se recogerán en un registro de metadatos en sustitución del valor literal, o junto a él.

El tercer aspecto contemplado son los *Conjuntos de elementos de Metadatos* o *Conjuntos de Elementos*, que se definen como las clases y atributos que se usan para describir ciertas entidades de interés. En la terminología de los datos enlazados, estos conjuntos de elementos se formalizan mediante esquemas RDF (*Resource Description Framework*)

o mediante ontologías OWL (*Web Ontology Language*) para los que se usa frecuentemente el término que los agrupa: *Vocabulario RDF*.

Por lo general, los conjuntos de elementos de metadatos no describen entidades bibliográficas, más bien proporcionan elementos para que se usen en la descripción de esas entidades, es decir son los elementos rellenables en un contexto leíble por máquina.

Sin duda, las tradiciones bibliotecarias se incorporan y trascienden con la vida que le otorga la tecnología. Sin embargo estas herramientas bibliotecológicas no contradicen el hecho irrefutable de que el Internet y la web han cambiado la vida de la población que tiene acceso a sus servicios y les ha otorgado la posibilidad de conocer todo tipo de información y a disponer de la cultura y la recreación en cualquier sitio.

La web incluye muchos entornos para compartir recursos. Algunos de ellos se utilizan para acciones educativas, publicar textos, páginas web, imágenes, audio, presentaciones multimedia, marcadores sociales; una gran variedad de situaciones, todas ellas indizadas con lenguaje natural o a través de títulos significativos de los contenidos, asignados por quienes comparten recursos, de alguna manera tienen el valor agregado que le otorga el ser pensante que es el hombre.

Un tipo más sofisticado de entorno tecnológico son las wikis en la que los usuarios pueden editar las páginas a través de navegadores web donde además crean y editan contenidos con herramientas al alcance de todos que permiten trabajar en colaboración y hacer sugerencias sobre las designaciones temáticas escogidas para incorporar a las páginas. Es decir, la tecnología necesita al hombre y lo incluye con la finalidad de hacer sistemas entendibles.

La wiki más conocida es la *Wikipedia* que se trata de un diccionario enciclopédico en la que cualquier persona puede escribir o modificar los textos. Entre las restricciones en el uso de las herramientas, se exigen títulos únicos para las entradas, a fin de poder ligar páginas wiki entre sí. Una entrada entre dobles corchetes de una página wiki la convierte en un enlace a otra página.

Este gran éxito de indización del conocimiento en una enciclopedia no está exento de reflejar las ideologías particulares que aquejan a los hombres. De acuerdo con la obra de Cesar Rendueles, profesor de filosofía de la Universidad Complutense de Madrid, la *Wikipedia* está

controlada por aquellos que generosamente dedican su tiempo regularmente a la edición de artículos. Es decir, se trata de una práctica social de conocimientos compartidos y dista bastante de un producto identitario de una inteligencia con un compromiso duradero de implicación política. (Rendueles, 2013)

En resumen, las wikis son intensamente colaborativas y esa es la razón de su éxito: su carácter interactivo, participativo y colaborativo (Godwin, 2003).

La intervención del razonamiento humano para seleccionar, organizar y transmitir, al igual que en las tradiciones bibliotecarias son una base sólida en el entorno actual de los servicios de información ofrecidos en la web, con independencia de una responsabilidad individual de peso intelectual sociocognitivo.

## FLUIDEZ E INTEROPERABILIDAD

Como se puede inferir del párrafo anterior, el valor agregado a la información se obtiene mediante la retroalimentación. Para discernir, vincular, seleccionar documentos, hacerlos visibles y con rapidez debe existir la participación e intervención humana.

La información circula en Internet con gran fluidez gracias a las técnicas modernas de comunicación y a las tecnologías que se van incorporando para interconectar la información disponible y accesible al instante. Cualquier persona que tenga acceso a una computadora disfruta la oportunidad de obtener información de todo tipo a través de un cierto control lingüístico que se ejerce en los sitios que la contienen.

El gran cambio en los últimos años en la tecnología permite la retroalimentación humana y el gran elemento incorporado en los sistemas de información tradicionales es la interoperabilidad.

La interoperabilidad en especial en la Bibliotecología es el elemento clave en el que se trabaja a través del movimiento que se explicó en el apartado anterior como los *Datos enlazados* (*linked data*) que es la integración de información de diferentes bases de datos cada vez más generalizada e internacional porque se incluyen opciones de búsqueda en gran cantidad de idiomas.

Los *datos enlazados* se incorporan a una *Central de Datos* (o *Data Hub*) que es el sitio en el cual se puede compartir información sobre paquetes de datos de cualquier tipo. La *central de datos* organiza los paquetes de datos en grupos que están a cargo de una comunidad. Los datos se encuentran en la *Nube de los Datos Abiertos Enlazados* (LOD, *Link Open Data Cloud*), donde también se encuentran los *Datos Enlazados Bibliotecarios* (LLD, *Library Link Data*) que incluyen tanto conjuntos de datos como vocabularios controlados.

Cada grupo tiene sus convenciones para describir los paquetes a incluir en él y prepara la documentación que explica el volumen de datos, recursos de ejemplo, métodos de acceso y los enlaces a otros paquetes de datos.

Seguir las convenciones de los dos grupos LOD y LLD garantiza que sus relaciones con otros paquetes están documentadas y que se contabilizarán como parte del corpus creciente de los datos enlazados.

Al tener los datos documentados de forma consistente es posible construir herramientas para alcanzar una mayor comprensión de su naturaleza y de la forma en la que encajan juntos. Además de ser interesante por sí mismo, este proceso es de importancia debido a que esta clase de conocimiento hace más fácil determinar cuál de los paquetes de datos es más apropiado para una determinada función; esto es, los datos resultan más fáciles de utilizar. (Informe del grupo *W3C Incubator*, 2011, s.p.)

La lectura del informe del grupo incubadora enlista recursos bibliotecarios enlazados: listas de encabezamientos de materia, vocabularios controlados y recursos generados para las bases de datos bibliográficas que se incorporan a la web con la finalidad de lograr significado y proseguir la marcha hacia la web semántica y la lista sigue creciendo día a día en los datos enlazados bibliotecarios que se van agregando. Sin estos datos para indizar información, difícilmente la web semántica podría generalizar el éxito hacia el sector que confía en las bibliotecas.



## LAS ONTOLOGÍAS

En un artículo publicado en la edición online de *ERCIM (The European Research Consortium for Informatics and Mathematics)*, que es el sitio europeo del *World Wide Web Consortium W3C*, en octubre de 2002, Tim Berners-Berners-Lee y Eric Molinero anticipan por segunda vez que la web alcanzaría su capacidad máxima cuando se convirtiese en un ambiente donde los datos pudiesen ser compartidos y procesados por las herramientas automatizadas y retroalimentados por la gente. Estaban refiriéndose a la web 3.0 o web semántica.

Para construir esta web semántica se necesita representar el conocimiento de forma que sea legible por las computadoras, esté consensuado y sea reutilizable. Las ontologías son la vía propuesta para representar este conocimiento.

La red semántica permite que la información no sólo esté ligada para su búsqueda a través de la red de forma automática, sino que también liga esta información a través de los metadatos y esto sirve para integrar y automatizar nuevas aplicaciones.

Algunos autores consideran que las ontologías serán el ladrillo de las próximas webs semánticas, ya que estas permiten relacionarse y usarse de una forma sencilla. Sin embargo los ladrillos aparecen a cuentagotas.

Las ontologías trabajan con personas, ideas, conceptos, objetos, instituciones y en general todo lo que en Bibliotecología consideramos entidades temáticas. Esos conjuntos de entidades temáticas en las ontologías son relacionados en una forma más detallada que en los tesauros tradicionales porque surgieron en el medio de cómputo y fueron trabajadas desde programas informáticos que facilitan su construcción y mantenimiento. Como indica Fernández Breis.

Las ontologías proporcionan un vocabulario de un área y definen niveles de formalismo, el conocimiento con ontologías se formaliza usando cinco tipos de componentes: clases, relaciones, funciones, axiomas e instancias. (Fernández Breis, 2008, p.41)

Por ahora, las ontologías ofrecidas y exitosas no son muchas. Entre ellas se encuentra FOAF, mencionada en líneas arriba de este trabajo que enlaza personas con personas y recursos web, de ella se obtiene buena respuesta. Si se observa la lista de ontologías disponibles en el documento de *Incubadora de Datos* es más corta, que la ofrecida por los medios bibliotecarios.

La información disponible en Internet crece en forma desmesurada, el número de herramientas de la web semántica aumenta cada día y los avances se presentan en situaciones aisladas como es el caso de los sistemas bibliotecarios. En contraste los buscadores siguen evolucionando a un ritmo acelerado. Los índices no responden en forma consistente a las inquietudes de los internautas que buscan información calificada y sobre todo con la oferta de intermediación semántica que promete la web semántica.

## INDIZACIÓN TEMÁTICA DESDE LA BIBLIOTECOLOGÍA

Las reglas para organizar, intercambiar y transferir la información, por medio de resúmenes y palabras significativas que representan los contenidos documentales son el dominio del bibliotecólogo. Ahora los modos de indizar y transmitir se han diversificado y el trabajo bibliotecológico al estar intrincado con la tecnología, parecería en este momento de la historia, ser dominado por ella.

Sin embargo, no es así, los grandes sistemas de información y las bibliotecas continúan con la selección, preparación, organización y transferencia de la información, respaldados con la tecnología ofrecida en cada etapa de avance de la misma. Es importante reconocer que el trabajo de indización en particular es lento y costoso porque involucra la acción humana de entender los conocimientos informados, resumirlos y representarlos a través de términos significativos que dirigen a los usuarios hacia los contenidos.

Cada vez es más necesario crear puentes a través de palabras significativas de los textos. Esta tarea por encontrar las palabras claves la compartimos con los editores, los comunicadores, periodistas y todos

los que participan en el ciclo de producción de conocimiento para transferir a la sociedad.

Por ejemplo, los editores pueden convencer a los autores para encontrar títulos atractivos que impulsen al lector a la compra de sus obras, a la hora de otorgarles una clasificación temática también buscarán el mejor ángulo que responda a sus intereses. Los periodistas escriben sobre acontecimientos en particular o entrevistan a autores y más bien se ocupan de lo trascendente en lo cotidiano del presente o del pasado, filtrado por su propia óptica.

Desde el punto de vista periodístico, para acaparar la atención del usuario es necesario encontrar la mejor expresión que represente un contenido y lo convierta en atractivo. El año anterior apareció una noticia con un pequeñísimo titular *Ronaldinbo al Querétaro*, que al mismo tiempo eran las palabras claves de la noticia. En sólo cuestión de minutos se dieron más de cinco mil menciones y 10 millones de impresiones de la expresión en las redes sociales y se consultó hasta en Alaska y Groenlandia. Sólo dos nombres conjuntados para darle visibilidad a la noticia.

El último ejemplo es claro desde el punto de vista periodístico porque juega con información global de entretenimiento, pero cuando se trata de noticias políticas o de la transmisión del resultado de una investigación policial, que a veces no ostentan la claridad suficiente o no respetan convenciones sociales o el discurso no contiene palabras claves políticamente correctas, las cosas se pueden complicar mucho.

La dificultad de escoger las palabras claves adecuadas en un texto o discurso para transferir información está avalada por múltiples casos de usos desafortunados o al contrario de éxito asegurado por un término atractivo.

La gran diferencia que ofrece la indización temática desde la Bibliotecología es que su interés está centrado en el usuario que busca adquirir un conocimiento concreto y su mecanismo de acercamiento a los sistemas de información bibliotecológicos se produce porque le ofrece documentación: seleccionada previamente con criterios de objetividad, publicados después de pasar la evaluación por pares y además organizada para acceder a ella con facilidad.

Los índices creados por el bibliotecólogo, además de ser claros, deben ser confiables del contenido que le anuncia al usuario y considerar el contexto desde el cual se produjo el texto y desde el cual busca la información el usuario.

La inmensa cantidad de datos que recibe en el día a día el usuario provoca cambios en su atención y obliga a los sistemas de información a ahondar en la comprensión de los textos para transmitir el contenido, asegurando al usuario que el tema asignado está tratado al interior del documento, expresa un conocimiento y no perderá tiempo si lo lee. El bibliotecario para expresar ese contenido utiliza los vocabularios controlados.

Los vocabularios controlados cuentan con el respaldo tecnológico del SKOS (acrónimo de *Simple Knowledge Organization System*), una iniciativa del W3C, que proporciona un modelo para representar la estructura básica y el contenido de esquemas conceptuales como listas de encabezamientos de materia, taxonomías, esquemas de clasificación, tesauros y cualquier tipo de vocabulario controlado. La web semántica supone una oportunidad para utilizar vocabularios controlados o tesauros en los procesos de descripción y recuperación de información con una evaluación permanente de los mismos.

El modelo SKOS tiene un planteamiento dirigido hacia una integración de los metadatos en los documentos para conseguir una retroalimentación lingüística que analice mediante tecnologías propias de la informática, el uso de los términos, con la finalidad de actualizar los sistemas de recuperación.

El medio bibliotecario publicó las normas *Information and documentation. Thesauri and interoperability with other vocabularies, ISO 25964-1* de 2011 y la *ISO 25964-2* de 2013 para actualizar el manejo tecnológico.

La norma *ISO 25964* en su primera parte cubre todos los aspectos de la elaboración de tesauros con una nueva visión de sinónimos, monolingües o multilingües, un desarrollo más puntual de las jerarquías relacionadas por medio del uso de verbos y el fomento de la interoperabilidad en la web, que incluye un modelo de datos y un esquema XML para el intercambio de datos.

La parte dos publicada en el 2013 tiene como principal objetivo fomentar la recuperación de información de alta calidad a través de los

recursos en la web que se han indizado con diferentes vocabularios. En él se explica cómo configurar las asignaciones entre los conceptos en estos vocabularios, y otras formas de uso complementario. Es decir, con la primera parte de la norma se establece el requisito de preparar para interoperar el tesoro a elaborar y la segunda parte trata de la interoperabilidad del tesoro con otros lenguajes controlados y entidades bibliográficas de recursos informativos.

Esta norma responde a la necesidad de contar en la web con una visión en la que se comparta información intercambiable en los servicios de etiquetado que se ofrecen a partir de la representación de los contenidos documentales. Esta mayor amplitud se puede mostrar por ejemplo, al contar con las diferentes definiciones que se pueden dar de un solo término en la especialidad de aplicación del vocabulario controlado, que las jerarquías se desglosen por las propiedades o atributos de las meronimias (casi todos los términos genéricos mantienen relaciones *todo-parte* en sus términos específicos).

No deben existir límites para los términos equivalentes, ya que se debe tratar de recoger todos los posibles usos para agrupar temas que puedan ser de interés en las búsquedas realizadas en los sistemas.

Además es imprescindible una mayor riqueza en las relaciones asociativas, típicas en las listas de encabezamientos y tesauros documentales, pero que no se reflejaban en los antiguos catálogos de referencias.

Dos años antes de la aparición de la segunda parte de las normas en la que se contemplan la interoperabilidad de los vocabularios controlados habían visto la luz los resultados del Grupo *Incubadora de datos enlazados* por lo que fueron incorporados.

Así, la organización tradicional se enfrenta a nuevos requisitos relativos a las representaciones estructuradas y flexibles de conocimiento que apoyan el descubrimiento de nuevos saberes, a partir de fuentes dispares y la inferencia sobre las muchas piezas de información. Esto requiere también la interoperabilidad sintáctica y semántica que le otorga el uso de la tecnología de cómputo. Sin duda la computación y la informática han influenciado en el cambio que se opera en la indización temática por el respaldo que ofrece, sin embargo el trabajo de base es el tradicional, al que se le incorporan además, los avances de tipo humanístico.

Como se ha visto, el valor agregado a la información no se va a obtener en todos los casos googleando, el bibliotecólogo tiene que discernir, vincular, seleccionar documentos, hacerlos visibles y con rapidez, a través del uso de las tecnologías apropiadas que se deberán conocer y aprender a manipular.

Quizás los cambios y las tendencias humanísticas más importantes tienen que ver con los modelos desde los que se realiza la indización temática al aplicarse la corriente socio-cognitiva, recogida por Hjørland (2002) tomada de la comunicación y la psicología. En esta corriente existe una relación mayor con los términos en los documentos hacia el usuario y un comportamiento más cercano al medio socio-cultural donde se desarrolla. Esto no quiere decir que no se incluyan los usos terminológicos en otros países, pero siempre privilegiando el ambiente socio-cultural donde se va a utilizar.

## CONSIDERACIONES FINALES

El trabajo intelectual requiere de una tecnología que debe ser formal y computarizada para ser eficiente. Pero lo que la gente sabe, lo que les gustaría saber, y lo que otros han aprendido y escrito sobre un tema, resiste el tratamiento mecánico. Si fuera de otro modo, la educación y la organización del conocimiento podrían reducirse a la informática.

La respuesta de la bibliotecología desde la organización de la información es brindar la seguridad que los temas asignados están reflejados en el contenido de los textos indizados y además trabajar para el futuro. En definitiva, la descripción de los textos hablados, escritos o en imágenes pretende transmitir una producción hecha pública en el pasado o el presente hacia el futuro.

Los índices y sistemas de categorización como tesauros, ontologías, taxonomías, folksonomías utilizan un lenguaje y por lo tanto, están en conflicto con la necesidad de mantener etiquetas inequívocas y estables para responder a los sistemas de organización. Es probable, que no siempre respondan a la consulta de un usuario, pero las probabilidades son mayores por el control y congruencia de un lenguaje seleccionado para contextualizar la información.

Los *datos enlazados bibliotecarios* en general están pensados en inglés y aunque los sistemas de traducción son muy sofisticados para algunas herramientas como el caso del *MeSH* o el Sistema de Clasificación Dewey, el Sistema de Clasificación Decimal Universal, *Europeana* o *Agrovoc*, es necesario trabajar desde el español para desarrollar datos enlazados en un contexto cercano, significativo y representativo de los productos intelectuales generados en español.

Los usuarios interactivos participan en la web indizando los contenidos que van incorporando para asegurarse que se recuperen. El éxito de la web es la interacción humana en la organización de la información que ofrece al asignar títulos, temas, autores y relaciones entre los elementos. Es probable que los índices finales sean imperfectos, pero es información etiquetada, en cierto modo relacionada con la organización bibliotecaria.

El saber es cultura, viva y cambiante, la tecnología es cultural también, pero inanimada y estable dentro de ciertos parámetros. Es decir, se puede estar seguro que la tensión entre los beneficios de la técnica y las limitaciones que genera la fijación del conocimiento estarán siempre presentes. La respuesta es continuar desarrollando sistemas de información interactivos con el ser humano que sirvan para recuperar en el futuro, mientras se espera el avance sólido e improbable de un sistema semántico automatizado.

## BIBLIOGRAFÍA

- Cumbre Iberoamericana de Jefes de Estado y de Gobierno. Reuniones Ministeriales Sectoriales (2010) *Bases para una estrategia iberoamericana de interoperabilidad. Documento marco iberoamericano de interoperabilidad*. Recuperado en mayo de 2014 [http://www.segib.org/publicaciones/files/2010/12/libroreuniones2010\\_castellano.pdf](http://www.segib.org/publicaciones/files/2010/12/libroreuniones2010_castellano.pdf)

David, P. A. y D. Foray (2002) “Una introducción a la economía del saber.” *Revista Internacional de Ciencias Sociales*, no. 171. Recuperado diciembre 2014 en: <http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/SHS/pdf/171-fulltext171spa.pdf>

Fernández Breis, J. T. (2008). *Un entorno de integración de ontologías para el desarrollo de sistemas de gestión de conocimiento*. (Tesis Doctoral) Universidad de Murcia. <https://digitum.um.es/xmlui/handle/10201/185>

Godwing, B. (2003). “Blogs and Wikis: Environment for On-line Collaboration.” *Journal of Language Learning & Technology*, 7(2), p.12-16.

Hjørland, B. (2002.) “Epistemology and the Socio-Cognitive Perspective in Information Science.” *Journal of the American Society for Information Science and Technology*, 53 (4), 257-270.

International Standard ISO 25964-1:2011. *Information and documentation – Thesauri and Interoperability with other vocabularies. Part 1. Thesauri for information retrieval*. Geneva: ISO.

International Standard ISO 25964-2:2013. *Information and documentation – Thesauri and Interoperability with other vocabularies. Part 2. Interoperability with other vocabularies*. Geneva: ISO.



*Library Linked Data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets.* W3C Incubator Group Report 25 October 2011. Recuperado mayo 2014 en: <http://www.w3.org/2005/Incubator/lld/XGR-lld-vocabdataset-20111025/>

Rendueles, C. (2013) *Sociofobia el cambio político en la era de la utopía digital*. Madrid: Capitán Swing, 200 p.

Savater, F. (2012) *El epicentro de la cultura es Internet*. Recuperado mayo 2014 en: <http://www.eluniversal.com.mx/notas/844837.html>

———, (2014) *Internet no educa; hay que educarnos para la red*. Recuperado mayo 2014 en: <http://www.tic.unam.mx/noticias/2014/10/2014-10-29-ciudadanos-en-la-red.html>

Stutt, A. and Motta, E. (2004). "Semantic Learning Webs. Journal of Interactive Media in Education." (10). *Special Issue on the Educational Semantic Web*. Recuperado mayo 2014 en: [ [www.jime.open.ac.uk/2004/10](http://www.jime.open.ac.uk/2004/10)]