

Algunas consideraciones sobre la evaluación de tesauros educativos y su inserción en plataformas virtuales educativas

CATALINA NAUMIS PEÑA
Universidad Nacional Autónoma de México

INTRODUCCIÓN

Dentro de las plataformas virtuales educativas juegan un papel relevante la organización de los contenidos educativos. El objetivo de este trabajo es proponer una metodología para evaluar tesauros y/o tomar la decisión de construir uno nuevo o adaptar alguno preexistente en un marco tecnológico que lo integre. La metodología es válida para tesauros sobre cualquier temática, pero en este trabajo se aplicará sólo a contenidos educativos, dada la necesidad de organización documental que exige la actual propuesta educativa a distancia. Para representar los programas de educación a distancia diseñados para contenidos educativos registrados sobre soporte digital y/o multimedia en los sistemas de información es necesario contar con un lenguaje documental que establezca los términos de indización del sistema educativo que habrá de soportar.

Los sistemas educativos a distancia necesitan compartir la información en forma segura, precisa y eficiente. Antes de abordar el problema de la evaluación de tesauros hay que estudiar el marco tecnológico en el que se insertan los tesauros, y las normas que se usan en los

sitios de educación a distancia para recuperar información de contenidos educativos.

NORMAS PARA RECUPERAR INFORMACIÓN DE CONTENIDOS EDUCATIVOS

La educación a distancia es mediada informáticamente para manipular la información que conviene a cada programa educativo en particular. Los sistemas de educación a distancia combinan los recursos educativos con las actividades pedagógicas de las personas, en sus correspondientes roles (Griffiths, Blat, García, Sayago, 2005: 6) De los elementos mencionados en este trabajo interesan los objetos de aprendizaje relacionados con los contenidos temáticos educativos, y su indización y recuperación. En el medio informático los contenidos son la esencia, más los metadatos (Jong, 2003: 9) por lo que es imprescindible aclarar que los contenidos que observaremos aquí son los contenidos temáticos educativos, y no, por ejemplo, los informáticos.

Los metadatos se utilizan para localizar, identificar, seleccionar y tener acceso a los objetos de aprendizaje. Los metadatos también documentan cómo se comporta ese objeto, cuál es su función, uso y relaciones con otros objetos, y cuáles sus características de visibilidad. Los estándares para intercambiar datos ayudan a definir las interfases con la finalidad de que los diferentes esquemas de metadatos puedan transferirse entre los sistemas de información educativos.

“Cualquiera que sea la tecnología que se use, debe permitir que el intercambio de información conserve el significado y la estructura original de los metadatos. Estas normas ofrecen una representación neutral de los metadatos y su orden estructural. No tienen que ver con la semántica subyacente, sino con proporcionar un modo común, legible por las máquinas, de transferir a lo largo de las redes, sistemas y plataformas los elementos que se han definido. Los estándares sobre intercambio de datos proporcionan así los interfases de mapeo entre el nivel de definición y el nivel tecnológico de los sistemas de información” (Jong, 2003: 16).

En general, los sistemas de información educativa están utilizando el lenguaje XML (Extensible Mark-up Language), que distingue entre forma y contenido (a diferencia del HTML, Hyper Text Mark-up Language), para traducir cualquier esquema de metadatos a un formato de representación común, con la finalidad de transferirlos a través de la web. El RDF (Resource Description Framework) funciona como una estructura global para intercambiar metadatos y proporciona el trabajo de base para otras normas. El RDF ha definido un modelo de metadatos de alto nivel y una sintaxis que se expresa en XML. El modelo RDF presenta tres tipos de objeto:

- un recurso; cualquier cosa que pueda tener una dirección URL (Uniform Resource Locator),
- una propiedad: un recurso con un nombre que puede usarse como una propiedad (ej. autor, título) y
- una declaración: consiste en la combinación de un recurso, una propiedad y un valor.

Si bien en el modelo RDF se manejan objetos, las relaciones entre ellos no son suficientes para responder con toda satisfacción las preguntas que podrían hacerse al sistema, y llegar así a los objetos de aprendizaje, por lo que el RDF se queda en niveles superiores de la estructura semántica y se han propuesto otras especificaciones basadas también en “programación objeto” en las que la búsqueda informativa escala a niveles interiores del contenido. Entre las propuestas se destacan la del IEEE (*Institute of Electrical and Electronics Engineers*) Learning Technology Standards Committee Learning Objects Metadata Working Group, la del Dublin Core Metadata Initiative, la del Global Learning Consortium que propuso el IMS Resource Metadata, y la del Advance Distributed Learning/Sharable Content Object Reference Model. Algunos de los involucrados en el desarrollo de estas normas comentan que:

“[...]sirven para desarrollar metadatos interoperables recíprocos para soporte tecnológico del aprendizaje, educación y herramientas de capacitación (Hodgins, 2001: 1), [...]tener una aproximación común a los metadatos educativos es crucial para promover la producción y

adopción de metadatos educativos. Esto, es lo primero en turno para avanzar en el camino largo de la infraestructura para educación abierta” (Duval, 2001: 2).

Pero las plataformas virtuales mencionadas evolucionan y por ejemplo Dublin Core, que comenzó proporcionando metadatos bibliográficos para describir recursos de texto, ha introducido ciertas extensiones para describir además recursos multimedia y audiovisuales que cubren aspectos específicos del audio y del video.

El modelo de metadatos debe conectar lógicamente las entidades, relaciones y atributos de metadatos a la propia esencia (digital) (Jon, 2003: 9). El modelo tendrá que enlazar los elementos informativos a la descripción de un todo que los comprende y definir el nivel de acceso dentro de los contenidos. Otra de las características de la norma utilizada será que pueda enlazar información textual, sonora e imágenes fijas y en movimiento.

LA INDIZACIÓN DE LOS CONTENIDOS TEMÁTICOS EDUCATIVOS

Como se desprende del apartado anterior los contenidos educativos temáticos en el ámbito de la informática son “objetos de aprendizaje” organizados en plataformas virtuales educativas. Además del modelo de interfase para recorrer y transmitir una idea general sobre la ubicación y características de los contenidos temáticos está el problema de la representación lingüística de esos contenidos.

La representación lingüística de los contenidos educativos es necesaria tanto en documentos digitales textuales como en multimedia. En los primeros la riqueza de las expresiones las hace confusas cuando se las consulta en busca de conceptos más que de palabras, y para no repetir una expresión en un mismo texto educativo, los usuarios de la plataforma educativa pueden usar palabras que no sean las de uso habitual. El proceso se hace largo y complejo cuando el usuario no localiza el tema y debe hacer el ejercicio de buscar todas las posibles formas en que un mismo concepto pueda haber sido expresado...

“Cuando se busca información acerca de conceptos que puedan ser expresados de diferentes formas es más efectivo usar directorios Web clasificados. Un control dirigido para sinónimos y homónimos, contextualizando los términos del índice colocándolos en una estructura jerárquica” (Mai, 2004: 92).

Los descriptores de un tesoro son otro tipo de metadatos específicos del contenido y se utilizan principalmente como herramientas de indización. En los documentos multimedia digitales la representación lingüística es básica para localizar las imágenes que ilustrarán un contenido educativo que consiste en ceros y unos. Las imágenes educativas tienen que transmitir exactamente el tema que se quiere enseñar para no dispersar al educando y aunque se dice que “una imagen transmite más que mil palabras”, la experiencia acumulada de los profesionales de la información permite asegurar que una palabra puede llevarnos a una imagen y no siempre sucede al revés.

En el entorno tecnológico actual se pueden usar muchos puntos de acceso

“[...]la computadora es capaz de almacenar y organizar textos de manera tal que pueda recuperarlos y de forma que el conocimiento pueda ser organizado en archivos digitales en el Internet usando una variedad de puntos de acceso potenciales” (Andersen, 2002: 37).

Este hecho que es una gran ventaja, como nos explica Andersen, se convierte en un problema cuando se recupera demasiada información. La discreción no es una virtud sólo del ser humano, es también deseable en los sistemas de información.

La plataforma virtual de enseñanza no soluciona la recuperación conceptual de información si se guía únicamente por las palabras del texto. El soporte de un tesoro es básico para designar los temas de forma consistente. Sin embargo la creación de un tesoro es costosa y difícil, por lo cual siempre se recurre a analizar las propuestas disponibles en el mercado para solucionar la organización del conocimiento.

LA EVALUACIÓN DE TESAUROS

Para aterrizar en un problema práctico, se evalúan tesauros existentes en el campo educativo con la finalidad de determinar la consistencia y armonía que tienen en relación con la propuesta de un nuevo tesoro. En la búsqueda de una estructura clasificatoria tipo tesoro para indizar aspectos educativos, la segunda tarea a desarrollar fue analizar los tesauros educativos existentes y establecer una metodología para aplicarle a cada uno de ellos.

Dadas las características del trabajo que se presenta no es posible ahondar en los detalles, pero se resume la metodología aplicada y el resultado de la evaluación. La metodología propuesta es enfocada desde cinco puntos de vista: análisis de la presentación, análisis de la consulta, consistencia, análisis del contenido y estructura terminológica y semántica.

Los cinco puntos de vista son importantes. La *presentación* porque si bien un tesoro contiene en su parte fundamental un corpus alfabético de descriptores e índices con entradas diferentes a las del corpus alfabético, para indizar y recuperar información, antes de esas estructuras debe explicarle al usuario una serie de aspectos que le servirán para decidir si el tesoro es útil en el sistema de información en el que se pretende insertar, y aprovechar en todo su potencial la herramienta que se le ofrece. La presentación tiene que ver no sólo con la mayor o menor facilidad de manejo, sino con la explicación que dan los autores sobre el nivel de estructuración conceptual del campo temático de aplicación.

La *consulta* define el modo de acceder al tesoro: si se está en línea, en disco compacto o en impreso, y contiene la fecha de actualización de los términos incluidos. La *consistencia* implica la observación de los siguientes elementos: que haya relaciones recíprocas, que los términos en iguales circunstancias tengan igual forma, que se den aclaraciones sobre la aplicación de los términos, y que haya normalización del género y traducción a otro idioma. En el *análisis de contenido* se rescata la relación entre el número de descriptores y no descriptores, los niveles de jerarquización y los tipos de contenidos que se incluyen. La *estructura terminológica y semántica* del

tesauro supone el análisis de la división temática del tesauro; es decir, la disección y organización de la terminología.

La aplicación de una metodología para revisar tesauros existentes, no sólo se hace con el objeto de descartar la posibilidad de utilizar uno preexistente y de no construir otro nuevo, sino también con el ánimo de cotejar aspectos parciales problemáticos del tesauro que se deberá generar, además de comparar los criterios empleados en el diseño de las categorías más generales y, por supuesto, de aprovechar aquellos elementos útiles para el tesauro que se va a desarrollar.

Así pues, si bien “[...]un tesauro para organizar contenidos educativos tiene que ser el fiel reflejo del sistema educativo al cual va a servir” (Naumis, 2002: 9), se procedió a analizar los siguientes tesauros educativos, por las ventajas metodológicas y prácticas ya explicadas:

- *Tesauro de la educación UNESCO: OIE* / Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. – 5ª ed. – París: UNESCO, 1992. – 144 p.
- *Thesauro de la UNESCO: lista estructurada de descriptores para la indización y la recuperación bibliográficas en las esferas de la educación, la ciencia, las ciencias sociales y humanas, la cultura, la comunicación y la información* / Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. – París: UNESCO, 1995. – 705 p.
<http://databases.unesco.org/thessp/> Fecha de consulta octubre 4 del 2005
- *Tesauro UNBIS: lista trilingüe (español, inglés, francés) de términos usados como encabezamientos de materia en el análisis de documentos y publicaciones relacionadas con los programas y actividades de las Naciones Unidas* / Biblioteca Dag Hammarskjöld. – 3ª ed. – Nueva York: Naciones Unidas, 1996. – 740 p.
[http://unhq-appspub-01.un.org/LIB/DHLUNBISThesaurus.nsf/\\$\\$\\$searchs?OpenForm](http://unhq-appspub-01.un.org/LIB/DHLUNBISThesaurus.nsf/$$$searchs?OpenForm)
- *Tesauro SPINES: un vocabulario controlado y estructurado para el tratamiento de información sobre ciencia y tecnología para el desarrollo* / Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. Instituto de Información y

Documentación en Ciencia y Tecnología. Madrid: UNESCO. ICYT, 1988 - 2 v.

http://unesdoc.unesco.org/ulis/cgi-bin/ulis.pl?database=ged&req=2&by=3&sc1=1&look=new&sc2=1&text_p=inc&text=SPINES&submit=GO

- *Macrotesauro para el procesamiento de la información relativa al desarrollo económico y social* / actualizado por Anne Di Lauro y Alice Watson. - 5ª ed. - París: Naciones Unidas. OCDE, 1998. - 427 p.

<http://shr.aaas.org/thesaurus/help.html>
[Fecha de consulta octubre 6, 2005]

- *Thesaurus of ERIC Education Research Information Center.* - http://www.eric.ed.gov:80/ERICWebPortal/Home.portal?_nfpb=true&_pageLabel=Thesaurus&_nfls=false
[Fecha de consulta: octubre 5, 2005].

La selección de dichos tesauros se basó fundamentalmente en el enfoque general que adoptan (es decir, tratan sobre educación), los aspectos concretos abordados y su actualidad (son tesauros recientes que se usan para incluir material audiovisual en los sistemas de información, y están centrados en el empleo de nuevos recursos didácticos).

El esquema empleado para analizar los tesauros se elaboró de acuerdo con ciertos indicadores que se tomaron, adecuaron y complementaron entre ellos, conforme a los trabajos de Alvaro Bermejo(1989) y Michel Dauzats (1994), y se les agregaron indicadores propios. Sin embargo en el presente trabajo, que sólo constituye un avance y un resumen de la investigación que se lleva a cabo, se hará mayor énfasis en la evaluación terminológica y semántica del tema del tesoro por considerarlo el centro del planteamiento tesauro; los otros cuatro aspectos que se evalúan contribuyen a hacer una apreciación general del servicio que brinda el tesoro.

La revisión de tesauros para seleccionar la muestra permitió apreciar que la mayoría de ellos han sido elaborados por instituciones internacionales, entre las cuales es menor el número de tesauros cuyo radio de acción es nacional y cuyos ámbitos científicos son restringidos y ni siquiera están dedicados a apoyar un sistema de enseñanza.

Puede resumirse el campo que abarcan los tesauros analizados de la siguiente forma: desarrollo económico y social, educación en un contexto internacional, ciencia y tecnología para el desarrollo, programas de las Naciones Unidas, otro con áreas diversas (educación, ciencia, cultura, ciencias sociales y humanas, cultura, información y comunicación, política, leyes y economía, países y grupos humanos) y, finalmente, investigación educativa.

El tesaurus de la UNESCO realizado por la Oficina Internacional de Educación está dividido en ocho campos semánticos, y en principio la definición en facetas debiera estar centrada en temas educativos. Sin embargo, una vez consultados los contenidos de esos campos semánticos se comprueba que la estructura y los términos responden a proyectos y actividades de la UNESCO. Incluso la sección tercera está dedicada específicamente a términos relativos a la enseñanza; es decir, los conocimientos transmitidos mediante un proceso de enseñanza incluyen términos de carácter demasiado general para un sistema de información educativo. Lamentablemente, el siguiente Tesaurus de la UNESCO (organizado en microtesauros) tiene las mismas particularidades y la misma limitación que el anterior.

Aunque en el Tesaurus UNBIS, se parte de una estructura temática principal no todos los descriptores forman parte del orden jerárquico y el tesaurus no cumple con la estructura entrelazada que se requiere para darle mayor consistencia.

En cambio el Tesaurus SPINES está organizado en treinta y cuatro campos temáticos y todos los descriptores que lo integran pertenecen a uno de ellos, con lo cual constituye un tesaurus de contenidos educativos en ámbitos restringidos, así, el nivel educativo que representa es de nivel profesional hacia arriba. La limitación que se observa para funcionar en un sistema educativo de nivel elemental y medio es el lenguaje científico y la cobertura temática. En conclusión: puede ser adoptado como modelo para determinados ámbitos temáticos educativos usando sinónimos lingüísticos adecuados al nivel de los usuarios.

El Tesaurus de la OCDE es un excelente modelo para indizar los aspectos relativos al desarrollo económico y social porque cuenta con una estructura en diecinueve clases, con una profundidad de siete jerarquías.

ERIC constituye, sin duda, el tesoro sobre temas educativos más conocido a escala internacional, y se ha convertido en la columna vertebral que articula el sistema de información internacional sobre educación. Sin embargo, después de su análisis se observa que está dirigido a la investigación educativa y no a los contenidos educativos. Además, esta herramienta lingüística está escrita en inglés, por lo cual debería ser traducido al español antes de poder ser aprovechado. Como dice Grijelmo, “[...]el idioma constituye la expresión más fiel de cada pueblo, y por eso ningún otro idioma podrá definirnos” (Grijelmo, 2002: 283). Incidiendo en el mismo aspecto, Yule afirma que

“[...]la lengua propia nos da un sistema acabado de categorizar lo que percibimos, y una consecuencia de ello es que este modelo le lleva a uno a percibir el mundo que le rodea a partir únicamente de esas categorías”(Yule, 1998: 280).

Del análisis de los tesauros educativos mencionados anteriormente se extrae la conclusión de que ninguno de ellos responde realmente a las necesidades de indización de un sistema que contenga documentos multimedia de apoyo a un programa educativo de nivel elemental y medio en idioma español.

Posteriormente se realizó la segunda parte del análisis que se propuso desde el principio sobre la organización de contenidos educativos que están en plataformas educativas. Del estudio realizado sobre algunos sitios y cuya discusión será publicada con posterioridad se obtuvieron soluciones para proponer un esquema tipo taxonómico, es decir que incluya el detalle jerárquico del sistema educativo, como base de inserción de un tesoro relativo a contenidos educativos.

Mientras tanto se menciona una propuesta de tipo tecnológico “ROSA: Repository of Objects with Semantic Access for e-Learning” que parte de la organización del propio sistema educativo. El primer conjunto de la categoría principal comienza con los diferentes programas de estudio y los diferentes niveles educativos que son impartidos, en la categoría subsiguiente se incluyen los nombres de los diferentes cursos y las clases temáticas son los contenidos de cada uno de los cursos. Si bien la organización de la información utiliza la misma premisa institucional, no fue posible revisar los contenidos para evaluarlos porque

no aparece aplicada en la Web, aunque existen artículos donde se explica su organización.

Cualquier tesoro que funcione en un ámbito educativo debería partir de una categoría principal que refleje las temáticas abordadas en los diferentes planes de estudio y luego ahondar en cada uno de ellos relacionándolos entre sí.

CONSIDERACIONES FINALES.

Más allá del resultado de la evaluación practicada es necesario reconocer que la metodología empleada permitió el acercamiento y entendimiento de la organización tesauro y brindó elementos de ayuda para crear un tesoro, pero casi se podría asegurar que es difícil encontrar un tesoro que mantenga una estructura de conocimiento útil para cualquier otro caso. Cada espacio educativo tiene sus objetivos, nivel de profundización, lenguaje utilizado, idioma y las características organizacionales de la institución que lo sustenta. La evaluación terminológica y semántica del tema del tesoro es el aspecto más significativo, los otros contribuyen a su apreciación.

En el análisis de la literatura bibliotecológica sobre el tema y partiendo de la premisa que el tesoro es un esquema clasificatorio, se encontraron elementos que afianzan la conclusión de la dificultad que existe para reutilizar un tesoro surgido de un contexto en otro:

“Un esquema de clasificación es un camino potencial para describir un dominio particular o el universo del conocimiento. Para crear un sistema de clasificación para una compañía, organización, biblioteca o cualquier otro centro de información en particular, se necesita comenzar con el estudio del discurso y de las actividades que se desarrollan en una organización o dominio. Se necesita aprender el lenguaje usado en la comunidad, desde que debe reflejar y responder al discurso particular de la comunidad. Una clasificación no es algo que pueda ser creado para una organización por una autoridad epistémica, la clasificación debe ser el resultado de la organización. La clasificación es la tipificación del lenguaje en una organización” (Mai, 2004: 46).

La dificultad estriba en la propia dinámica social, cultural y económica de las instituciones en las que surge la necesidad de organizar los contenidos que generan “[...]el concepto de organización del conocimiento está en interacción con y deriva de la organización social del conocimiento” (Andersen, 2002: 37).

De acuerdo con las conclusiones previas contrastadas con la educación elemental y media, México se propone organizar el esquema clasificatorio del tesoro para contenidos educativos multimedia con diferentes tipos de metadatos: los primeros que reflejen los diferentes niveles educativos que se imparten, otros para los diferentes cursos comprendidos en cada nivel educativo y los últimos para los contenidos educativos propiamente dichos, que serán desarrollados en la interrelación temática acostumbrada en un tesoro. La parte tecnológica deberá cuidar las relaciones entre los tres tipos de metadatos para permitir la opción de recuperar cualquier contenido educativo con la finalidad de abrir la opción de usarlos en programas de difusión o extensión de la cultura.

OBRAS CONSULTADAS

- Alvaro, C., Villagrà, A. & Sorli Rojo, A. (1989), “Desarrollo de lenguajes documentales formalizados en lengua española: II. Evaluación de los tesauros disponibles en lengua española”, en *Revista Española de Documentación Científica*. – Vol. 12, no. 3, p. 283-305.
- Andersen, J. (2002), “Communication Technologies and the Concept of Knowledge Organization”, en *Knowledge Organization*. Vol 29, no. 1, p. 29-39
- Duval, E. (2001), “Metadata Standards Leaders IEEE and DCMI Collaborate to Design Future Metadata Architecture for Web-based Learning, Education and Training”, en *Standards IEEE* Retrieved February 15, 2006.
<http://standards.ieee.org/announcements/metaarch.html>

Griffiths, D. Blat, J. García, R. & Sayago, S (2005) *La aportación de IMS Learning Design a la creación de recursos pedagógicos reutilizables*. Universitat Pompeu Fabra. Retrieved february 17, 2006.

<http://www.upf.edu>

Grijelmo, A. (2002) *Defensa apasionada del idioma español*. México: Taurus

Hodgins, W (2001), "Metadata Standards Leaders IEEE and DCMI Collaborate to Design Future Metadata Architecture for Web-based Learning, Education and Training", en *Standards IEEE* Retrieved February 15, 2006.

<http://standards.ieee.org/announcements/metaarch.html>

Jong, A. (2003) *Los metadatos en el entorno de la producción audiovisual* / traducción de J. Andréz. 2ª ed. México: Radio Educación.

Le Thesaurus de l' image: etude des langages documentaires pour l' audiovisuel (1994) / sous la direction de Michel Dauzats. París: ADBS Editions.

Mai, J. E. (2004), "Classification in Context: Relativity, Reality and Representation", en *Knowledge Organization*. Vol 31, no. 1, p. 39-48

Naumis, C. (2001) *El macrotesauro mexicano para contenidos educativos: manual para uso y retroalimentación*. México: CUIB (documento interno)

Repository of Objects with Semantic Access for e-Learning

Yule, G. (1998) *El lenguaje* / Traduc. N. Bel Rafecas. Madrid: Cambridge University Press.