

# INSTRUMENTOS DE UTILIDAD EN EL ANÁLISIS DEL CONTENIDO PARA LA ORGANIZACIÓN DE DOCUMENTOS DIGITALES EDUCATIVOS

Miguel Ángel MARZAL GARCÍA-QUISMONDO, mmarzal@bib.uc3m.es

María Jesús COLMENERO RUIZ, mcolmene@bib.uc3m.es

Aurora CUEVAS CERVERÓ, accerver@bib.uc3m.es

Universidad Carlos III de Madrid (Getafe, España). Departamento de Biblioteconomía y Documentación

## Resumen

Las políticas de la Unión Europea para la sociedad de la información están impulsando un modelo educativo basado en adquisición de competencias y el aprendizaje a lo largo de la vida. La eficacia del aprendizaje en este modelo educativo descansa en una educación digital, en la que los recursos digitales estructuran, organizan y representan su contenido acorde con sus potencialidades educativas. La Biblioteconomía y Documentación puede contribuir a su consecución mediante la actualización de sus instrumentos a partir de su diseño conceptual en el análisis documental del contenido. Se analizan tesauros, mapas conceptuales y ontologías, que constituyen diversos sistemas de representación del conocimiento procedentes de diferentes disciplinas, y se realiza una comparativa en cuanto a su eficacia para la organización de contenidos en educación para el espacio digital.

## Palabras clave

competencias en el uso de la información, recursos electrónicos en la educación, educación a distancia, análisis del contenido, mapas conceptuales, ontologías, tesauros

## 1 INTRODUCCIÓN

Las políticas educativas de la Unión Europea para la sociedad de la información, consecuencia del impacto derivado de la difusión y las funcionalidades de las tecnologías de la información y comunicación, se han constituido en verdaderas *políticas de información para la educación*. El apoyo a la *e-Education*, ya desde la acción e-Europe 2002,<sup>1</sup> se ha visto incrementado exponencialmente en el ulterior desarrollo de las actuaciones europeas encaminadas a impulsar la referida sociedad de la información.

1. <[http://europa.eu.int/information\\_society/eeurope/2002/index\\_en.htm](http://europa.eu.int/information_society/eeurope/2002/index_en.htm)>

Dos son los aspectos clave sobre los que giran estas políticas. El primero, el cambio hacia un modelo educativo donde el elemento activo del aprendizaje es el alumno, basado en las teorías del aprendizaje significativo, en el «aprender a aprender». Este modelo ha de desenvolverse en un espacio distinto al tradicional, un nuevo espacio educativo que trasciende el espacio del aula, extendiéndose en el espacio digital, creando entornos colaborativos, dinámicos y asíncronos (las «aulas expandidas»).

El segundo aspecto se centra en potenciar e impulsar el *aprendizaje a lo largo de la vida*, sustentado en la necesidad de dotar a los ciudadanos de las competencias adecuadas para que puedan adaptarse a los cambios y alcanzar la capacitación adecuada durante toda su vida profesional, más allá de la alcanzada tras el periodo inicial de formación.

Ambos elementos de cambio pivotan sobre el mismo eje, un modelo educativo basado en la adquisición de competencias, en contraposición a la capacitación basada en un conjunto cerrado de conocimientos y el dominio de destrezas y habilidades. La eficacia del aprendizaje en este modelo educativo, construido sobre competencias, desarrollado en un nuevo espacio, donde el elemento central es el alumno y proyectado hacia el aprendizaje continuo, reposa en una *educación digital*, en la que los recursos digitales se estructuran, representen y organicen conforme con sus potencialidades educativas.

La adaptación de la educación universitaria a este modelo a través del Espacio Europeo de Enseñanza Superior surgido a partir de la Declaración de Bolonia,<sup>2</sup> está impulsando a su vez la necesaria transformación de las bibliotecas universitarias en centros de recursos para el aprendizaje y la investigación (CRAI).<sup>3</sup>

En este escenario, las Ciencias de la Documentación pueden contribuir con dos aportaciones relevantes: un espacio educativo organizado [mediante el control, identificación, búsqueda y recuperación de recursos informativos para Educación, el soporte de las comunidades virtuales y la edición y gestión de materiales educativos en los CRAI y CRA<sup>4</sup> (MARZAL 2005)] y unos instrumentos emergentes de organización de contenidos, a partir de su diseño conceptual en el análisis documental de contenido (ADC). Un análisis de estos últimos es el objeto de nuestra contribución.

2. Los recientes Reales Decretos que regulan los estudios de grado y postgrado en nuestro país son el reflejo de la puesta en marcha de las políticas europeas de educación para la plena incorporación de sus ciudadanos a la Sociedad de la Información (Reales Decretos 55/2005 de Grado y 56/2005 de Postgrado, de 21 de enero, publicados en BOE nº 21 de 25 de enero de 2005 <<http://www.boe.es/g/es/boe/dias/2005-01-25/seccion1.php#00003>>).
3. Rebiun, organismo que aglutina a las bibliotecas universitarias españolas, está impulsando la investigación y la transformación de éstas en CRAI, con el fin de proporcionar el apoyo adecuado al modelo educativo universitario europeo. Pueden consultarse documentos al respecto en la web de Rebiun y, en especial, en el enlace correspondiente a sus II Jornadas Rebiun <[http://bibliotecnica.upc.es/rebiun/nova/jornadas/segundas\\_jornadas\\_rebiun/intro.htm](http://bibliotecnica.upc.es/rebiun/nova/jornadas/segundas_jornadas_rebiun/intro.htm)>.
4. Centros de Recursos para el Aprendizaje. Suponen el equivalente a los CRAI en la educación no universitaria, como transformación de la biblioteca escolar.

## 2 Condicionantes de los recursos digitales para su eficacia educativa

El nuevo modelo educativo delineado requiere unos nuevos instrumentos capaces de soportar el proceso inferencial hacia el conocimiento de los datos y bloques informativos de la red, con características distintivas frente a los tradicionales. Así, el diseño y la aplicación de instrumentos que permitan alcanzar la adecuada efectividad didáctica de los recursos digitales se ve condicionada fundamentalmente por:

- Las potencialidades tecnológicas de los recursos digitales: el hipertexto modifica los modos de comprensión y aprendizaje de la información, estableciendo una relación con el texto habilitada para una lectura digital asociativa, politextual y selectiva. Si bien el hipertexto imita la disposición del texto tradicional, la forma en que el texto llega al lector determina su comprensión,<sup>5</sup> haciéndose útil, por tanto, una *cibersemiótica*.
- La funcionalidad educativa de los recursos digitales: ésta posee una doble dimensión; de un lado, la *interacción* que permite la comunicación interpersonal de los usuarios en el «espacio» del recurso y, por tanto, el *aprendizaje colaborativo*. De otro, la *interactividad*, relación recíproca entre usuario y recurso, orientado a un aprendizaje desde la *inteligencia emocional*, de modo que el recurso se adapte a las destrezas del usuario.

En la educación digital estos dos condicionantes deben orientarse hacia la optimización de la eficacia didáctica de la *navegación*, elemento sustancial al recurso digital, por su capacidad de sustentar un proceso intelectual complejo. Un buen sistema de navegación debe interconectarse con una adecuada estructuración de la información, dado que los usuarios no reproducen la forma física en la que la información está estructurada, sino las relaciones conceptuales subyacentes para construir las representaciones mentales de su contenido (SCHAWN 2002). Los aspectos cognitivos involucrados en un proceso de aprendizaje, pues, precisan fórmulas de navegación basadas en conceptos acomodados a las asociaciones semánticas horizontales.

## 3 EL ANÁLISIS DOCUMENTAL DEL CONTENIDO FRENTE A LOS DOCUMENTOS DIGITALES

La acción tradicional del análisis documental del contenido (ADC), al quedar enfrentada a la singularidad de los documentos digitales, se ve impulsada a una transformación que la habilite para ser aplicada adecuadamente en este medio.

5. Los investigadores del campo de la cultura escrita, como Roger Chartier, han puesto de relieve que los cambios producidos en la presentación de los textos han influido tanto en los modos de leerlos como en su comprensión.

Tres son los condicionantes principales que, en el proceso que lleva desde la información hacia el conocimiento, obligan a que este cambio se produzca (MOREIRO 2001).

El primero es consecuencia de la primacía del usuario en el espacio digital, pues ejerce la función de lector, tomando sus propias decisiones en cuanto al camino a recorrer por el hipertexto, determinadas por su interés. En un ambiente de aprendizaje, la principal necesidad de este usuario-lector es llegar de la información al conocimiento, haciéndolo a través de la selección. La información se produce en paquetes, y es el usuario el que debe escoger cuáles utilizará, de entre la multiplicidad de opciones que se le presenten. Así, los usuarios seleccionan y generan su propia colección documental digital, que irá siendo más personal según su nivel cognitivo.

El ADC, por tanto, debe convertir sus lenguajes documentales en instrumentos más cercanos al lenguaje de los usuarios, no al de los documentos, favoreciendo la manipulación correcta de los sistemas informativos. Darles una orientación «centrada en el usuario», con la flexibilidad suficiente para responder a las demandas de éste con la suficiente rapidez y eficacia, de forma que primen las competencias del usuario en cuanto a la selección y evaluación de los recursos pertinentes para sus necesidades informativas (FERNÁNDEZ-MOLINA 2002).

El segundo condicionante viene dado por la dilución de la noción de «materia» en los contenidos. La tradicional separación académica en disciplinas ha ido transformándose en un entramado de relaciones múltiples entre ellas, llevando a la multidisciplinariedad e interdisciplinariedad. En el medio digital, esta situación se hace especialmente patente, por lo que el ADC debe tratarlo como conceptos o centros de interés, en línea con las necesidades del usuario en cuanto a navegación y búsqueda de información. Las representaciones basadas en conceptos así obtenidos mejoran el aprendizaje cuando el material se organiza en torno a ellas (SMITH 2002).

El tercer y último condicionante proviene de la necesaria alfabetización en información (*information literacy*) o alfabetización informacional. En 1989 la American Library Association (ALA) la entendió como medio de formación permanente, mediante la que se dota de competencia para hallar información, organizar su conocimiento y usarlo para aprender. En 1998 aparecieron las primeras normas sobre alfabetización en información, también dentro de la ALA, que fueron el precedente de las *Information Literacy Standards for Higher Education* publicadas por la *Association of College and Research Libraries* (ACRL), perteneciente a ALA, en 2000<sup>6</sup> y las *Information Literacy Standards* del *Council of Australian University Librarians* en 2001, actualizadas en febrero de 2004.<sup>7</sup> La alfabetización en información se entiende como el nuevo modo de aprender en entornos tecnológicos, una vez adquiridas las competencias necesarias por

6. <<http://www.ala.org/ala/acrl/acrlstandards/informationliteracycompetencystandards.htm>>

7. <<http://www.caul.edu.au/info-literacy/InfoLiteracyFramework2003spanish.doc>>

parte del educando para poder abordarlo en óptimas condiciones (BAW-DEN 2002). El ADC debe acudir a organizar los contenidos desde la perspectiva de generación de bibliotecas virtuales «educativas» (MARZAL 2003a, 2003b), esfuerzo que se une al de los centros de recursos para el aprendizaje y la investigación (CRAI) en los ámbitos de educación superior ya citados.

Los sistemas de clasificación precoordinados, como la clasificación decimal universal (CDU) o los encabezamientos de materia, fundamentalmente taxonómicos, son propios de una organización jerárquica del conocimiento que refleja las disciplinas tradicionales. La difuminación paulatina de las fronteras que las delimitaban está provocando la revisión de este tipo de instrumentos, por las dificultades encontradas para acomodarlos sin cambios a esta nueva situación (MOREIRO 2001). El instrumento del ADC tradicional más adecuado para este espacio digital, por su capacidad asociativa y relacional, es el tesoro. También éste se encuentra constreñido en parte y está asistiendo a una evolución de sus postulados teóricos que permitan moldearlos a las exigencias actuales.

Analizaremos a continuación tesauros, mapas conceptuales y ontologías, que constituyen diversos sistemas de representación del conocimiento procedentes de diferentes disciplinas pero que poseen características adecuadas para la representación del conocimiento con una estructura asociativa, teniendo como punto de mira su eficacia en la organización de contenidos en Educación.

### 3.1 Tesauros

La automatización y el desarrollo de la web han ido evidenciando, por un lado, la adaptación de los tesauros a nuevos usos, integrándose, en conjunción con el lenguaje natural, en los sistemas de gestión de la información (SGI) como bases de conocimiento para mejorar la recuperación de la información<sup>8</sup> y en los sistemas de búsqueda en la red y, por otro, sus posibilidades hipertextuales a través de las relaciones para servir como elementos de navegación.

Paralelamente, se configura como necesidad la revisión de la estructura de los tesauros con el fin de aprovechar las posibilidades que la tecnología ofrece y adaptarlos al medio digital, lo que ha confluído con las críticas expresadas sobre la debilidad o indefinición de algunos conceptos de la norma.

Las modificaciones posibles se extienden a distintos aspectos, que incluyen tanto el vocabulario como las relaciones. Una de las propuestas más completas es la que hace Antonio García Gutiérrez (GARCÍA GUTIÉRREZ 1998) denominando *lenguaje epistemográfico* a un nuevo tipo de lenguaje documental que supone la «*evolución teórico-metodológica de los te-*

8. Las bases de conocimiento constituyen una de las partes estructurales de los sistemas automáticos «basados en conocimientos» o «sistemas expertos» desarrollados por la inteligencia artificial.

*sauros*», haciéndolo avanzar hacia un lenguaje mucho más asociativo y adaptado a las posibilidades tecnológicas actuales. Sin embargo, no es el único autor que reclama una evolución del tesauruso tradicional.

En cuanto al vocabulario, se apuesta por la inclusión de otras formas léxicas distintas del sustantivo o palabras sustantivadas prohibidas en la normativa para su inclusión como descriptores, como es el caso de verbos (MARZAL 2002), adjetivos o adverbios. Se incide además en la necesidad de definiciones para los términos y de reglas más concretas en relación con la actuación ante los términos compuestos, cuya utilización ha ido en aumento, en particular los del tipo sintagmático, para incrementar la precisión en la recuperación (KURAMOTO 2002).

Las relaciones de equivalencia pueden quedar reducidas agrupando todos los términos equivalentes en torno al concepto, pudiéndose acceder a éste desde cualquiera de ellos, con lo que se convierte en un auténtico nodo de una red semántica o mapa conceptual (MOREIRO 2001) por lo que se ha llegado a hablar de «tesauros conceptuales». Es esta también la intención que guía las propuestas de tesauros de usuarios, que pretenden un diseño centrado en el usuario acercando el lenguaje utilizado por éste al del indizador, adoptando premisas de la psicología cognitiva. Las relaciones jerárquicas, aun bien definidas en la norma, no son todas en puridad relaciones ontológicas (relación es-un o *is-a*), por lo que dan lugar a problemas de integración y reutilización, incluso en otros instrumentos como las redes semánticas y las ontologías, pasando algunas a la consideración de relaciones asociativas (DOERR 2001).

Las mayores críticas a la normativa actual se centran en la infrautilización, falta de definición y sistematización de las relaciones asociativas. En un esfuerzo por superar estas deficiencias se han realizado algunos intentos de interés en relación con la sistematización de las relaciones que existen entre conceptos,<sup>9</sup> planteándose las ventajas que la inclusión de una tipificación de éstas en los tesauros tendría para la recuperación de la información (TUDHOPE 2001).

Los tesauros de construcción reciente suelen incorporar parte, al menos, de las modificaciones que se están demandando. Como ejemplo se pueden citar el *Getty Art and Architecture Thesaurus* (AAT),<sup>10</sup> el *US National Library of Medicine Unified Medical Language System* (UMLS),<sup>11</sup> metatesauruso que reúne más de 50 vocabularios, AGROVOC,<sup>12</sup> tesauruso desarrollado por la FAO (*Food and Agriculture Organization*) de Naciones Unidas, o *Laurin*.<sup>13</sup>

Todo este movimiento teórico se ha entrelazado además con la irrupción de la web semántica, lo que ha provocado que la norma deba empezar a revisarse, no sólo desde el punto de vista conceptual sino también

9. Uno de ellos se desarrolló en el seno de uno de los comités de la ALA (American Library Association) <[www.ala.org/ala/alctscontent/catalogingsection/catcommittees/subjectanalysis/subjectrelations/finalreport.htm](http://www.ala.org/ala/alctscontent/catalogingsection/catcommittees/subjectanalysis/subjectrelations/finalreport.htm)>. [Consultado: 20/01/2005].

10. <[http://www.getty.edu/research/conducting\\_research/vocabularies/aat/](http://www.getty.edu/research/conducting_research/vocabularies/aat/)>

11. <<http://www.nlm.nih.gov/research/umls/>>

12. <<http://www.fao.org/agrovoc/bar.htm>>

13. <<http://laurin.uibk.ac.at/>>

desde un punto de vista tecnológico e integrador.<sup>14</sup> En este contexto, el avance hacia la interoperabilidad de las bases de conocimiento ha desembocado en el impulso de las ontologías, como veremos, necesitándose la adaptación de los tesauros a estas nuevas exigencias y a los lenguajes desarrollados para ella (SOERGEL 2004). Ya se han elaborado distintas propuestas en una variedad de formatos de intercambio.<sup>15</sup>

Aunque aún es pronto para saber cómo serán los cambios que se avecinan y el papel que los tesauros tendrán en este nuevo entorno, lo que queda patente es su consideración como instrumentos útiles para organización, búsqueda y navegación en el espacio digital.

### 3.2 Mapas conceptuales

Los mapas conceptuales son un instrumento proveniente del campo de la didáctica, desarrollado por Joseph D. Novak y su equipo en la Universidad de Cornell (Estados Unidos) como resultado de sus investigaciones sobre el proceso de aprendizaje infantil de las ciencias en los años setenta. Él mismo los define como «*herramientas para organizar y representar conocimiento*».<sup>16</sup>

Para su desarrollo, se basaron en las teorías sobre la psicología del aprendizaje de del psicólogo cognitivo David Ausubel, enunciadas en los años 60, con el que Novak colaboró posteriormente. Ausubel desarrolló la teoría de la asimilación, centrada en el concepto de aprendizaje significativo, que establece que el aprendizaje se produce mediante la asimilación de nuevos conocimientos, conceptos y proposiciones, que adquieren significado al incorporarse a la estructura cognitiva previa del individuo. Esta estructura cognitiva está formada, según Ausubel, por un conceptos y proposiciones organizados de forma jerárquica, representando su conocimiento y experiencias.

El aprendizaje significativo es un proceso dinámico, basado en la interacción entre los conceptos ya existentes y los nuevos, sirviendo aquéllos como base para la elaboración de significados nuevos, y la creación de relaciones significativas entre ellos. Para que la información se asimile debe ser conceptualmente clara y coherente con el conocimiento previo del aprendiz. Así, la intención de Novak al desarrollar los mapas

14. La norma americana ANSI/NISO Z39.19-1993, *Guidelines for the Construction, Format, and Management of Monolingual Thesauri* fue revisada en 1998. No obstante, se observó la necesidad de su adaptación al entorno digital lo que llevó a la NISO, organismo encargado de la normalización en EEUU, en 2003 a iniciar un nuevo proceso de revisión, actualmente en curso <<http://www.niso.org/committees/MT-info.html>>. Por otra parte, el Comité Técnico TC46/SC 9 de la ISO está revisando la norma correspondiente a los tesauros multilingües ISO 5964:1985.

15. La iniciativa SWAD-Europe (W3C *Semantic Web Avanced Development for Europe*), proyecto financiado por la Unión Europea (UE) como parte del programa *Information Society Technologies* (IST-2001-34732), incluye un grupo de trabajo dedicado a los tesauros, SWAD-Europe Thesaurus Activity. Su página Web recoge los avances del proyecto y una página de enlaces que sistematiza los esfuerzos realizados en este área <<http://www.w3.org/2001/sw/Europe/reports/thes/>>

16. <<http://cmap.coginst.uwf.edu/info/>>

como instrumento metodológico de ayuda al aprendizaje significativo es que sirva como ayuda y refuerzo al educando, exteriorizando las relaciones de conocimiento que éste posee, la estructura del conocimiento y sus procesos de construcción (metacognición), en definitiva, aprendiendo a aprender.

El objetivo de un mapa conceptual es representar relaciones entre conceptos en forma de proposiciones con el fin de hacerlas explícitas. Joseph D. Novak define concepto como «una regularidad percibida en sucesos u objetos o registros de sucesos u objetos, designado por una etiqueta», etiqueta que suele ser una palabra, y proposición como «una frase acerca de cierto objeto o suceso en el universo, que ocurre de forma natural o artificial. Las proposiciones contienen dos o más conceptos conectados con otras palabras que forman una frase coherente», indicando que también son denominadas «unidades semánticas».

Conceptos y relaciones se representan en un gráfico bidimensional, donde los conceptos, incluidos generalmente en círculos o cajas, forman nodos y las relaciones entre ellos son indicadas mediante una línea de conexión etiquetada con palabras o frases que especifican el tipo de relación que enlaza los dos conceptos. Así, la secuencia de conceptos y enlaces forman proposiciones o frases simples mediante las palabras usadas en los nodos y los enlaces (Figura 1). El mapa conceptual más simple estaría formado por dos nodos conectados por un enlace representando una frase sencilla. Por convención, los enlaces se leen de arriba hacia abajo a menos que incluyan una punta de flecha. También incluyen enlaces horizontales entre conceptos de dominios diferentes del mapa conceptual.

Novak defiende una estructuración en forma jerárquica de los mapas, con los conceptos más generales en la parte superior y los más específicos en orden descendente hacia la parte inferior, lo que para otros autores no es estrictamente necesario. En cualquier caso, no implican secuencia, temporalidad, direccionalidad o jerarquías de poder.

En la elaboración de mapas conceptuales, a diferencia de los tesauros, no se realiza ningún control del vocabulario ni normalización (MOREIRO 2002), no existiendo tampoco normas o reglas para realizar las asociaciones. El resultado final estará determinado por la visión que de ese dominio del conocimiento posea el autor del mapa, pues de él únicamente depende la elección tanto de los conceptos como de los enlaces semánticos, lo que motiva que su nivel de complejidad sea muy variable y que no exista un mapa «correcto».

Aunque el mayor uso de los mapas conceptuales se ha realizado en el ámbito educativo (como herramienta de aprendizaje, método evaluativo del proceso de aprendizaje en sí mismo y útiles para el desarrollo del *currículo*), de este campo se ha extendido a otros de variada índole, adoptando principalmente sus capacidades para entornos cooperativos.

Al igual que otras áreas, la automatización primero, y la web después, han llevado los mapas conceptuales al medio digital, derivando de su realización manual a la automatizada mediante programas específicos. Estos aportan ventajas para su construcción y mantenimiento, mejoran su



Las bases teórico-metodológicas de los mapas conceptuales que, a diferencia de los tesauros, no habían sido objeto de revisión salvo en lo referente a su estructura jerárquica, empiezan ahora a ser revisadas. Aunque los mapas conceptuales son apropiados para representar el conocimiento estático, el análisis de los conceptos y las proposiciones lleva a distinguir entre relaciones estáticas y dinámicas (SAFAYENI s.d.), proponiendo el uso de un nuevo tipo de mapa complementario, *Cyclic Concept Map*, para la adecuada representación del conocimiento de las relaciones dinámicas o funcionales entre conceptos.

### 3.3 Ontologías

A principios de los años 90, en el seno de la comunidad científica de la Inteligencia Artificial (IA) comienza a perfilarse una nueva área de interés centrada en el concepto «ontología» a partir del *Knowledge-Sharing Effort*.<sup>19</sup> Este proyecto se centraba en remediar los problemas de interoperabilidad entre las denominadas bases de conocimiento buscando una solución que permitiera compartir y reutilizar el conocimiento almacenado por dichos sistemas y su aplicación en redes. Construir las suponía un esfuerzo y coste considerable y el trabajo realizado no se podía, en muchos casos, utilizar de nuevo.

Consensuar una definición del término propició un proceso de debate durante casi toda la década que, si bien parece cerrado, su conclusión se circunscribe a la IA, sin dar una solución completa a los demás campos científicos en los cuales se fue extendiendo su uso y donde se suele utilizar con connotaciones distintas, sin permitir precisar los límites que tiene con otros términos relacionados, especialmente el de modelo conceptual.

La definición dada por Gruber<sup>20</sup> es una de las más conocidas y aceptadas: «una ontología es una especificación explícita de una conceptualización», a pesar de algunos intentos de clarificación terminológica.

El acuerdo general se centra únicamente en los componentes básicos de una ontología: el vocabulario específico de un dominio con sus correspondientes definiciones, las cuales describen los conceptos, las relaciones que se dan entre ellos y las reglas que combinan ambos. Estas definiciones, cuya intención es evitar la ambigüedad, pueden especificarse con diferentes lenguajes, en un rango que abarca desde la forma más simple mediante una jerarquía, hasta las más complejas mediante la lógica formal, que añade axiomas para limitar los significados posibles y permitir el razonamiento.

En el contexto de la web semántica el interés de las ontologías gira alrededor de tres características: el ser un vocabulario compartido, con-

19. El *Knowledge-Sharing Effort*, fue promovido por la *Air Force Office of Scientific Research* (AFOSR), la *Defense Advanced Research Projects Agency* (DARPA), la *Corporation for National Research Initiatives* (CNRI), y la *National Science Foundation* (NSF). Puede consultarse información sobre las actividades y avances del proyecto en <<http://www-ksl.stanford.edu/knowledge-sharing/index.html>>.

20. <<http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>>.

sensuado en una comunidad de interés y reutilizable. De esta forma, pueden construirse grandes ontologías a menor costo mediante el ensamblaje de componentes ya existentes y, mediante un escalado de ontologías a distintos niveles, puede integrarse información de diferentes dominios.

Aunque se han propuesto clasificaciones de ontologías basadas en distintos criterios (GUARINO 1997; LASSILA 2001) las que se centran en el grado de formalidad con el que se especifica el vocabulario recogido nos resultan de especial interés puesto que a mayor formalidad la semántica que incorporan se incrementa. Más que en categorías discretas, las distintas ontologías se pueden representar a lo largo de una línea continua donde el grado de definición varía de un extremo a otro (Figura 2). Siguiendo la línea de izquierda a derecha el número de términos puede disminuir, su ambigüedad decrece y aumenta su significado, su semántica. La doble línea, arbitraria, indicaría el punto a partir del cual, según sus autores, la mayoría de los investigadores entienden que es una ontología. Sin embargo, en la literatura es frecuente que todos los tipos expuestos sean asimilados a una ontología.<sup>21</sup> Las diferencias estriban en su potencial expresivo según su nivel de abstracción y axiomatización, lo que proporciona una escala que permite una construcción ontológica por capas o niveles.

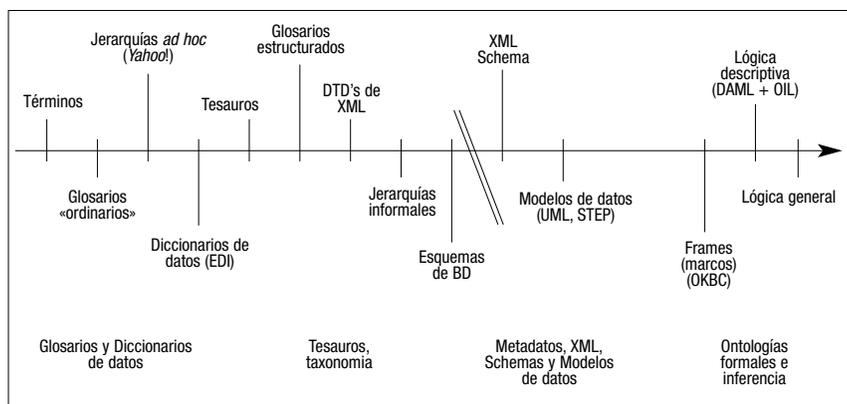


FIGURA 2. Tipos de ontologías según Gruninger y Uschold (2002)

Este escalado en capas obliga a la interoperabilidad entre ontologías, la cual no es tarea fácil. Dado que un concepto particular puede estar representado de múltiples formas según la ontología que lo recoja, es necesario establecer relaciones entre los elementos de distintas ontologías o mapeo, sin que se produzcan inconsistencias a lo largo de la ontología resultante.

21. Las alusiones al formato de metadatos Dublin Core, la jerarquía de *Yahoo*, los tesauros o *Wordnet* como ejemplos de ontologías son constantes en la literatura consultada.

## 4 COMPARATIVA ENTRE LOS INSTRUMENTOS ANALIZADOS

Un examen comparativo los instrumentos analizados nos lleva a destacar sus semejanzas y potencialidades, que recaen fundamentalmente en la aportación de elementos de asociatividad entre conceptos, acordes al interés que nos ocupa, o lo que es igual, la búsqueda de herramientas documentales que permitan la organización de los recursos educativos en el medio digital adaptados a las características del proceso de aprendizaje significativo.

Dos son los lugares comunes a todos ellos: los conceptos y las relaciones entre ellos. Con distintas denominaciones terminológicas y reglas diferentes para su uso, un ejercicio de abstracción los iguala en el plano teórico y conceptual. Esta similitud los ha llevado a un plano de interrelación al fundirse desde campos diferentes en un espacio común, el espacio digital. Espacio que a su vez se organiza sobre el hipertexto, cuya estructura basada en nodos y enlaces se superpone igualmente a la misma representación.

Los tesauros, que recogen la experiencia en el campo documental, y los mapas conceptuales, que sintetizan el de la educación, se incorporan a la web como instrumentos de representación del conocimiento, como ontologías de bajo nivel que constituirían los primeros escalones en la web semántica. Permiten revelar la estructura del conocimiento subyacente y, mediante su utilización conjunta, acercarse terminológicamente al usuario. Así mismo, permiten dar soporte al multilingüismo, factores, todos ellos, imprescindibles para su aplicación en los CRAI.

Los conceptos son representados en todos ellos mediante palabras extraídas del lenguaje natural. En mapas conceptuales y ontologías, a diferencia de los tesauros, no se realiza control sobre el vocabulario. Sin embargo, en las ontologías se intenta elegir términos representativos de los conceptos, estableciendo relaciones y axiomas con el fin de acotar su significado de forma más rígida que en los tesauros. Todos ellos, además, incluyen relaciones o enlaces entre conceptos, con distintos sistemas de utilización, grados de sistematización y denominaciones.

Las relaciones de equivalencia de los tesauros proveen un mecanismo para acercar todas las «etiquetas» posibles de un concepto al lenguaje del usuario, tal como suele estar representado en los mapas conceptuales. Este tipo de relación está ausente como tal en mapas conceptuales y ontologías, aunque se utiliza un mecanismo semejante cuando se integran en sistemas automatizados mediante su combinación con otras herramientas de tipo lingüístico.

La estructura jerárquica impuesta al esquema de representación del conocimiento del dominio es más laxa en los mapas conceptuales, donde su necesidad está puesta en entredicho por algunos autores y no está sujeta a reglas de construcción. Tanto en tesauros como en ontologías supone una parte fundamental, establecida a partir de relaciones jerárquicas que son definidas en forma similar y adolecen de los mismos problemas de imprecisión en ambos casos, permitiendo la inferencia a partir de mecanismos de herencia.

Las relaciones asociativas, por último, suponen el mayor reto en cualquiera de ellos. Los tesauros, aún teniendo definidas algunos tipos, dejan

abierta la puerta a la inclusión de otras. Los mapas conceptuales admiten cualquier tipo de relación asociativa que el autor quiera expresar y en las ontologías, donde se especifican como propiedades, dependen del sistema lógico elegido para definirlos.

Una diferencia fundamental de los tesauros en relación con los otros dos instrumentos es la ausencia de especificación del tipo de relación establecida. No poseen más descripción que uno de los tres grandes tipos de relación, lo que limita su expresividad semántica. Como ya vimos, paliar estas dificultades está siendo uno de los intereses de la investigación más reciente, con la intención de convertirlos en herramientas más potentes en el ámbito automatizado. Además, tanto mapas conceptuales como ontologías utilizan principalmente verbos para especificar el tipo de relación establecida siendo equiparables proposiciones a propiedades.

## 5 CONCLUSIONES

Tesauros y mapas conceptuales están inmersos en un proceso adaptativo al medio automatizado que los ha llevado, en una primera etapa, hacia las bases de conocimiento y, en una segunda, hacia las ontologías, acompañándose a esta evolución del campo de la IA. Sin embargo, ambos instrumentos carecen de la suficiente formalización, hasta el momento, que permitan su utilización como ontologías en sistemas que requieran un gran poder descriptivo que permita el razonamiento.

Por otro lado, la aplicación a la web de las ontologías muy formalizadas no es tan sencilla como parece. Los investigadores procedentes de la IA se encuentran con un medio en el que sus técnicas tradicionales no pueden aplicarse con eficacia, precisamente por la falta de estructuración de la información que contiene y la dificultad de su implantación en un sistema descentralizado si el nivel de aprendizaje necesario para crear contenidos para este medio aumenta.

De esta forma, sistemas menos formales, pero estructurados, como los tesauros y los mapas conceptuales parecen tener una adecuada aplicación en la web. Hay muchos desarrollados, con óptimos resultados, que son utilizados como base para la construcción de ontologías de complejidad creciente y como primer escalón en dominios reducidos. El hecho de que muchos aporten una estructuración de los datos los hace idóneos para su aplicación en la Web, que es el espacio sobre el que se desenvuelve la *educación digital*. Así, además de existir la posibilidad de integración de mapas conceptuales y tesauros como primeros niveles del escalado de ontologías, ambos instrumentos permiten su utilización como apoyo a la comprensión del área de conocimiento que describen.

La utilización de un instrumento asociativo que permita un acercamiento a las ontologías para su incorporación adecuada a la web semántica, sirviendo además de puente entre éstas y los tesauros, que recogen la experiencia en el campo documental, y los mapas conceptuales, que sintetizan el de la educación, sería idóneo como punto intermedio. Esto evitaría una superposición de capas o niveles intermedios elevada, con

los correspondientes problemas de integración, a la vez que permitiría conservar la capacidad de revelar la estructura del conocimiento subyacente y su cercanía terminológica al usuario.

Pensamos que un instrumento idóneo para este engarce, tal y como se ha descrito, centrado en el documento educativo, es el modelo expresado en la norma ISO/IEC 13250: 2003 *Topic Maps* cuyas posibilidades en este campo estamos empezando a indagar. Este estándar se basó en la semántica implícita en los índices analíticos impresos para su adaptación al procesamiento automático. Aunque su aparición y evolución se produjo en círculos tecnológicos, alejados del documental, no deja de hundir sus raíces en éste. Tanto tesauros como mapas conceptuales pueden ser expresados utilizando esta norma para adaptarlos al medio digital.

## BIBLIOGRAFÍA CITADA

- (BAWDEN 2002) BAWDEN, D. «Revisión de los conceptos de alfabetización informacional y alfabetización digital» [recurso electrónico]. *Anales de documentación*, n. 5 (2002), p. 361-408. <<http://www.um.es/fccd/anales/ad05/ad0521.pdf>>. [Consulta: 16 febrero 2005].
- (DOERR 2001) DOERR, M. «Semantic Problems of Thesaurus Mapping» [recurso electrónico]. *Journal of Digital Information*, vol. 1, n. 8 (2001). <<http://jodi.ecs.soton.ac.uk/Articles/v01/i08/Doerr/>>. [Consulta: 17 febrero 2005].
- (FERNÁNDEZ-MOLINA 2002) FERNÁNDEZ-MOLINA, J. C.; GUIMARÃES, J. A. C. «Ethical Aspects of Knowledge Organization and Representation in the Digital Environment: Their Articulation in Professional Codes of Ethics». En: LÓPEZ-HUERTAS, M. J. (ed.). *Challenges in knowledge representation and Organization for the 21st century. Integration of knowledge across boundaries: proceedings of the seventh international ISKO conference, Granada (España) 10-13 Julio 2002*. (Alemania: Ergon, 2002), p. 487-499.
- (GARCÍA 1998) GARCÍA GUTIÉRREZ, Antonio *Principios del lenguaje epistemográfico: la representación del conocimiento sobre Patrimonio Histórico Andaluz*. Granada: Instituto Andaluz del Patrimonio Histórico, 1998.
- (GUARINO 1997) GUARINO, N. «Understanding, Building, and Using Ontologies [recurso electrónico]: A Commentary to *Using Explicit Ontologies in KBS Development*, by van Heijst, Schreiber, and Wielinga». *International Journal of Human and Computer Studies*, vol. 46, n. 2/3 (1997), p. 293-310. <<http://www.inf.ufrgs.br/gpesquisa/bdi/links/articles/ontologies/guarino97understanding.pdf>>. [Consulta: 13 enero 2005].
- (KURAMOTO 2002) KURAMOTO, H. «Sintagmas nominais: uma nova proposta para a recuperação de informação» [recurso electrónico]. *DataGramaZero-Revista de Ciência da Informação*, vol. 3, n. 1 (2002). <[http://www.dgz.org.br/actual/Art\\_03.htm](http://www.dgz.org.br/actual/Art_03.htm)>.
- (LASSILA 2001) LASSILA, O.; MCGUINNESS, D. L. *The Role of Frame-Based Representation on the Semantic Web. Knowledge Systems Laboratory Tech Report Number KSL-01-02. Stanford University*. [recurso electrónico]. <[http://www.ksl.stanford.edu/KSL\\_Abstracts/KSL-01-02.html](http://www.ksl.stanford.edu/KSL_Abstracts/KSL-01-02.html)>. [Consultado: 11/03/2005].
- (MARZAL 2002) MARZAL GARCÍA-QUISMONDO, M. A. et al. «Propuesta para la utilización de estructuras verbales a la recuperación y representación de la información». *SCIRE: representación y organización del conocimiento*. vol. 8, n. 1 (2002), p. 95-102.

- (MARZAL 2003a) MARZAL GARCÍA-QUISMONDO, M. A. *et al.* «Modelos instrumentales de la documentación para recuperar conocimiento: Meta-datos y Topic Maps para educación». En: SIGEF CONGRESO (10º: 2003: León). *Emergent solutions for the information and knowledge economy*. (2003).
- (MARZAL 2003b) MARZAL GARCÍA-QUISMONDO, M. A. *et al.* «El recurso educativo como documento: learning objects». En: JORNADAS ANDALUZAS DE DOCUMENTACIÓN ((3as: 2003: Sevilla). *Organizaciones electrónicas: Situación actual y perspectivas de la E-documentación, Sevilla 20, 21 y 22 de noviembre 2003*. (Sevilla: AAD, 2003), p. 67-69.
- (MARZAL 2005) MARZAL GARCÍA-QUISMONDO, M. A.; CUEVAS CERVERÓ, A.; COLMENERO RUIZ, M. J. «La biblioteca escolar como centro de recursos para el aprendizaje». En: *V Congreso Internacional Virtual De Educación, 7-27 de febrero, 2005*.
- (MOREIRO 2001) MOREIRO GONZÁLEZ, J. A.; MARZAL GARCÍA-QUISMONDO, M. A. «Modelos teóricos y elementos funcionales para el análisis de contenido documental: definición y tendencias» [recurso electrónico]. *Investigación bibliotecológica*, vol. 15, n. 31 (2001), p. 125-162. <<http://www.ejournal.unam.mx/iibiblio/vol15-31/IBI03107.pdf>>. [Consulta: 4 enero 2005].
- (MOREIRO 2002) MOREIRO GONZÁLEZ, J. A. *et al.* «Mapas conceptuales, topic maps y tesauros». [recurso electrónico]. En: *I Jornadas de Tratamiento y Recuperación de la Información, 4-5 julio, 2002, Valencia*. <<http://www.fundacion.uc3m.es/jotri2003/ponencias/mapas.pdf>>.
- (SAFAYENI s.d.) SAFAYENI, F.; DERBENTSEVA, N.; CAÑAS A. J. *Concept Maps: A Theoretical Note on Concepts and the Need for Cyclic Concept Maps*. [recurso electrónico]. <<http://cmap.ihmc.us/Publications/ResearchPapers/Cyclic%20Concept%20Maps.pdf>>. [Consultado: 13/02/2005].
- (SCHAWN 2002) SCHAWN FARRIS, J.; JONES, K. S.; ELGIN PETER, D. «Users' Schematic of Hypermedia: What Is so "Spatial" about a Website?». *Interacting with Computers*, vol. 14 (2002), p. 487-502.
- (SCHMITZ-ESSER 2003) SCHMITZ-ESSER, W. «Meaning, Understanding, and the Organization of Knowledge in a Multilingual World – New Tools for New Tasks: Ontologies» [recurso electrónico]. *Sprachwissenschaft – Computerlinguistik – Computational Linguistics*, n. 16 (2003), p. 149-171. <[http://www.schmitz-esser.de/pdf/1\\_1\\_9\\_3\\_asiscr.pdf](http://www.schmitz-esser.de/pdf/1_1_9_3_asiscr.pdf)>. [Consulta: 26 enero /2005].
- (SMITH 2002) SMITH, T. R.; ZENG, M. L.; ADEPT KNOWLEDGE ORGANIZATION TEAM. «Structured models of scientific concepts for organizing learning materials». En: LÓPEZ-HUERTAS, M. J. (ed.). *Challenges in knowledge representation and Organization for the 21st century. Integration of knowledge across boundaries: proceedings of the seventh international ISKO conference, Granada (España) 10-13 Julio 2002*. (Alemania: Ergon, 2002), p. 232-239.
- (SOERGEL 2004) SOERGEL, D. *et al.* «Reengineering Thesauri for New Applications: the AGROVOC Example» [recurso electrónico]. *Journal of Digital Information*, vol. 4, n. 4 (2004). <<http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Soergel/>>. [Consulta: 14/02/2005].
- (TUDHOPE 2001) TUDHOPE, D.; ALANI, H.; JONES, C. «Augmenting Thesaurus Relationships: Possibilities for Retrieval» [recurso electrónico]. *Journal of Digital Information*, vol. 1, n. 8 (2001). <<http://jodi.ecs.soton.ac.uk/Articles/v01/i08/Tudhope/>>. [Consulta: 9 febrero 2005].
- (VISSER 2002) VISSER, U.; SCHLIEDER, C. «Modelling with Ontologies» [recurso electrónico]. En: STUCKENSCHMIDT, H.; STUBKJAER, E.; SCHLIEDER, C. (eds.). *The Ontology and Modeling of Real Estate Transactions in European Jurisdictions*. (Aldershot: Ashgate, 2002), p. 115-130. <[http://www.tzi.de/buster/papers/cost\\_visserschlieder.pdf](http://www.tzi.de/buster/papers/cost_visserschlieder.pdf)>. [Consulta: 24/01/2005].