

## TESAUROS Y LENGUAJES CONTROLADOS EN INTERNET

*José Luis Laguens García\**

Centro de Documentación de CEOE. Madrid.

**Resumen:** Se examinan las características, ventajas e inconvenientes de los lenguajes documentales en formato electrónico cuyo contenido es accesible a través de Internet. Para ello, se describen algunos ejemplos de clasificaciones y tesauros de diverso tipo que pueden consultarse en la Red.

**Palabras clave:** Lenguajes documentales; tesauros; clasificaciones; documentos electrónicos; Internet.

**Title:** THESAURI AND CONTROLLED VOCABULARIES IN INTERNET.

**Abstract:** This paper examines the characteristics, vantages and disadvantages of thesauri and other kind of controlled vocabularies in electronic format availables in Internet. To illustrate this, the author expose some examples of classifications and thesaurus that is possible to locate on the Web.

**Keywords:** Documental languages; classifications; thesaurus; electronic objects; Internet.

### 1. INTRODUCCIÓN

La 'World Wide Web' de Internet se ha convertido en uno de los principales vehículos de información de todo tipo. En los sitios Web que circulan por la Red es posible encontrar información sobre prácticamente cualquier tema que se pueda imaginar en un formato electrónico que permite transmitir no sólo textos escritos sino también documentos gráficos y audiovisuales. Internet es una fuente de información que ya está demostrando un enorme potencial de desarrollo. La difusión de documentos multimedia a través una red telemática mundial es un hecho incuestionable muy presente en la vida cotidiana que a estas alturas del desarrollo de las tecnologías de la información y de las telecomunicaciones no supone ninguna novedad, sino una realidad palpable en los ámbitos doméstico, laboral, académico o científico.

La cantidad de usuarios que utilizan Internet de manera habitual sigue creciendo día a día en todos los países. Los usuarios privados o las organizaciones utilizan la Web con propósitos muy diversos: comunicación de mensajes, intercambio de información y conocimientos, realización de transacciones comerciales, publicidad de productos y servicios, difusión de iniciativas, ideas y opiniones, etc. Esta expansión de Internet ha propiciado la transmisión universal de documentos de manera instantánea y sin la mediación del correo postal tradicional. Para las comunidades científica, académica e incluso empresarial, uno de los aspectos más interesantes de la Web es la capacidad de difundir y obtener conocimientos a través de servicios electrónicos de información bibliográfica y repositorios de documentos electrónicos. Para poder acceder a su contenido y saber de qué manera se estructuran dentro de cada uno de ellos, tanto los documentos que los integran como sus referencias bibliográficas, siguen utilizándose las clasificaciones y lenguajes documenta-

---

\* [documentacion@ceoe.es](mailto:documentacion@ceoe.es)

les, herramientas que facilitan la organización de los recursos documentales por temas y materias. Con la edición electrónica es posible elaborar lenguajes documentales en formato digital y aprovechar las ventajas que permite este soporte.

Los lenguajes documentales, sean clasificaciones terminológicas o tesauros, siguen siendo instrumentos indispensables para estructurar la información y el conocimiento en sistemas organizados de almacenamiento y difusión de documentos, sobre todo si tenemos presente que en la catalogación actual se tiende a conceder una mayor relevancia a los descriptores como puntos de acceso al registro bibliográfico. Esto incrementa el valor de los tesauros y otros lenguajes documentales como instrumentos de indización y consulta.

En el presente trabajo trataremos de describir, a grandes rasgos, algunas de las características de los lenguajes documentales convertidos en objetos digitales que pueden encontrarse en Internet, examinando para ello algunos ejemplos.

## 2. DOCUMENTOS ELECTRÓNICOS

Antes de examinar la ventajas (y también los inconvenientes) de los lenguajes documentales en formato digital, habría que delimitar el concepto de documento electrónico. En líneas generales, podría definirse como todo aquel documento que ofrece su contenido codificado en dígitos binarios registrados en cualquier tipo de soporte informático, como disquetes magnéticos, discos duros de ordenador, discos digitales (CD, CD-ROM, DVD) o servidores de redes, y que necesita para su reproducción dispositivos informáticos. Respecto al documento impreso en papel, el documento electrónico aporta una nueva forma de suministrar y mostrar contenidos, así como una nueva manera de acceder ellos y de manejarlos.

Dentro de este tipo de documentos han cobrado enorme auge las llamadas ‘páginas Web’, que constituyen la base de la información que circula en Internet. Se trata de documentos asociados a un nodo de la Red y escritos en un lenguaje específico (normalmente HTML) que pueden reunir archivos multimedia con información en diferentes formatos (PDF, GIF, MP3, TIFF, etc.) y pequeñas aplicaciones que ejecutan acciones concretas.

A parte de la gran ventaja que aportan integrando textos, gráficos, imágenes en movimiento y sonidos ofrecen la posibilidad de organizar su contenido mediante una estructura hipertexto con enlaces que facilitan el desplazamiento rápido a lo largo de las diversas partes que conforman el documento. Otras ventajas son también:

- La facilidad con la que pueden editarse, publicarse y actualizarse.
- Su gran accesibilidad y capacidad de difusión.
- La posibilidad para el usuario final de imprimir o almacenar su contenido.
- La inclusión de enlaces (‘links’) que remiten a otros documentos, sitios o páginas Web sobre temas afines o de especial relevancia.
- Suelen contener también las direcciones de correo electrónico de sus autores, de manera que los usuarios pueden contactar e intercambiar sugerencias y opiniones con ellos.

Naturalmente, este medio también presenta algunos inconvenientes derivados de su fuerte dependencia tecnológica y de sus originales características:

- Debido a que en la Web es posible ‘colgar’ todo tipo de páginas de muy diversa procedencia y en muchas ocasiones de duración efímera, la calidad de los contenidos va-

ría enormemente, siendo obligado seleccionar con acierto aquellos documentos que más se aproximen a nuestras necesidades de información y procedan de fuentes solventes y fiables.

- Es necesario conocer la dirección URL ('Uniform Resource Locator') exacta del sitio o la página Web a la que se desea acceder o bien recurrir a un buscador o portal temático, que agrupa según un índice por temas las páginas que localiza. Esto plantea dos problemas, por un lado, las direcciones URL son a menudo muy largas y cambiantes, dificultando su memorización, y por otro, para que una página Web tenga repercusión entre la audiencia a la que va dirigida es preciso promocionarla en la Red y enlazarla con otros sitios, buscadores o portales de interés.
- La tecnología en la que se basan la mayor parte de los documentos Web que circulan en Internet es el lenguaje de marcas HTML, además de contener a veces aplicaciones programadas en lenguajes como Java o JavaScript, posibilitando ser leídos por cualquier plataforma de ordenador o navegador. Sin embargo, las sucesivas versiones de HTML, que incluyen cada vez mayor cantidad de elementos no reconocidos por todas las plataformas y navegadores, están complicando los estándares y por tanto la permanencia de un lenguaje universalmente legible por todas las máquinas conectadas a la Web. Para resolver este problema se están desarrollando otros lenguajes alternativos que superen estos inconvenientes, como es el caso del lenguaje XML, que se considera probable sucesor de HTML.
- Un problema relacionado también con los aspectos tecnológicos es el de la velocidad de acceso a la Red proporcionada por las infraestructuras y los proveedores de Internet. Al ser cada vez más numerosos los usuarios, la capacidad de transmisión de los sistemas tiende a saturarse con mayor facilidad, causando retardos en la descarga de documentos Web o impidiendo su acceso en determinados momentos.
- También hay que mencionar los problemas de seguridad y manipulación por manos ajenas al creador del documento Web original. Dado que un documento electrónico básicamente no es más que un código digital, éste es fácilmente modificable por cualquier usuario con mínimos conocimientos sobre elaboración de páginas Web. Por lo tanto, resulta muy difícil garantizar el respeto a los derechos de autor en los documentos Web. Ni siquiera protegerlos mediante elementos de programación y códigos de seguridad ofrece plenas garantías.
- Otro inconveniente, todavía no del todo resuelto, es el de la conservación a largo plazo de los documentos electrónicos. Como no sabemos exactamente cuánto tiempo pueden aguantar en un mismo soporte, se hace necesario planificar la migración periódica a otro soporte de refresco. Este problema se agrava con la constante evolución de los formatos de archivo, los programas y las tecnologías, lo que obliga a pensar en estrategias para migrar documentos electrónicos antiguos conservando la funcionalidad de los formatos o programas en los que se originaron y que han caído en desuso por resultar ya obsoletos.

### **3. LOS LENGUAJES DOCUMENTALES COMO DOCUMENTOS ELECTRÓNICOS**

Los lenguajes documentales son instrumentos ideados para ayudar a la sistematización de colecciones de documentos con el objeto de facilitar su identificación y posterior recu-

peración, y su uso está muy extendido en centros de documentación, bibliotecas y archivos.

Una parte importante del análisis documental consiste en la indización de los documentos procesados, basada en la asignación de uno o varios términos incluidos en un lenguaje documental (clasificación o tesauro) previamente elaborado, con el fin de permitir la representación temática de cada unidad documental y su inclusión, junto a otros datos relevantes del documento, en una base de datos, facilitando así su consulta y posterior recuperación física. Los lenguajes documentales son pues instrumentos de referencia fundamentales para las tareas documentales de indización y recuperación, y deben estar siempre a disposición de documentalistas y usuarios.

Hasta la expansión del ordenador como herramienta de trabajo en las organizaciones, los lenguajes documentales se elaboraban sólo en papel y eran el resultado de un largo y costoso proceso manual de recopilación y edición. El manejo de una terminología en soporte impreso es más incómodo que en formato electrónico: se puede perder o deteriorar y su actualización a menudo requiere volver a imprimir el documento completo o gran parte del mismo. Un lenguaje documental electrónico elimina estos inconvenientes: se puede manejar desde el propio ordenador al mismo tiempo que los documentalistas o los usuarios utilizan otros recursos electrónicos, admite actualizaciones constantes y al instante sin necesidad de imprimir el documento, con el consiguiente ahorro de papel, y hace más cómodas y rápidas las tareas de indizar o consultar, pudiendo tener abierto el archivo que lo contiene mientras trabajamos. Asimismo, al tratarse de un documento electrónico, un mismo tesauro puede ser utilizado en red por varios usuarios a la vez y puede almacenarse en tantas copias de seguridad como se desee.

A grandes rasgos, se distinguen dos tipos de lenguajes documentales: las clasificaciones y los tesauros. Los primeros son lenguajes precoordinados, puesto que coordinan, antes del almacenamiento de los documentos, los diferentes conceptos que conforman una materia. De esta clase son por ejemplo las clasificaciones o listas alfabéticas de materias utilizadas en bibliotecas. Prevén por lo general una única materia asignable a cada documento y resultan bastante útiles para la ordenación física de los documentos. Por su parte, los tesauros son lenguajes postcoordinados, puesto que yuxtaponen los conceptos para coordinarlos después de la memorización de los documentos. Tales conceptos están representados por descriptores, que constituyen los términos admitidos y normalizados para la indización y la búsqueda documental y pueden asociarse libremente entre sí en función de la profundidad concedida a la indización. Aquí radica uno de los puntos fuertes de los tesauros, ya que con ellos puede realizarse una descripción más exhaustiva del contenido temático de los documentos, aumentando así sus posibilidades de recuperación desde diversos puntos de vista.

Los tesauros también permiten controlar los sinónimos y los homógrafos, haciéndolos remitir a los términos normalizados (descriptores), que serán los únicos válidos para la indización y la recuperación documental. Asimismo, incluyen 'cuasisinónimos', es decir, términos que no son exactamente sinónimos de descriptores válidos sino conceptos afines que pueden remitir a éstos. También muestran las diferentes relaciones jerárquicas y asociativas existentes entre los descriptores, además de incluir en ocasiones notas de alcance para aclarar nociones o definir conceptos. El uso de tesauros permite expresar conceptos de un modo más complejo, resolviendo también dificultades sintácticas mediante términos compuestos y otros recursos.

Actualmente, las bibliotecas y centros de documentación utilizan Internet como un instrumento más para realizar sus tareas y servicios. De este modo, es posible crear una página Web con el objetivo de dar a conocer a los usuarios tales servicios, poner en red un sistema de consulta de catálogos o facilitar la comunicación directa con el usuario a través del correo electrónico. El acceso electrónico a colecciones de referencias bibliográficas y repositorios de documentos digitales agiliza notablemente la difusión de la información, evitando la visita presencial de los usuarios al centro de documentación. Estos pueden formular consultas desde un terminal remoto situado en sus casas o lugares de trabajo sin necesidad de desplazarse y de manera instantánea, y en muchos casos leer o descargar en el disco duro de su ordenador documentos a texto completo. No obstante, las búsquedas deberán estar asistidas por algún procedimiento de ayuda y por instrumentos de referencia, como tesauros, clasificaciones o listas de materias. De ahí la utilidad práctica que supone contar con versiones electrónicas de los mismos.

Los documentos Web son idóneos para representar lenguajes documentales que puedan ser colgados en la Red formando parte de un sistema de ayuda para la formulación de búsquedas. La posibilidad de añadir funciones de hipertexto y objetos multimedia (muy útiles por ejemplo para crear tesauros legibles para discapacitados) aporta grandes ventajas a los tesauros, mejorando su manejabilidad y su aspecto visual, además de permitir su acceso a múltiples usuarios a la vez. Su actualización puede realizarse de manera más rápida y cómoda, sin esperar a costosas ediciones en papel. Además, Internet facilita la comunicación vía correo electrónico entre los responsables del mantenimiento del tesoro y los usuarios o expertos que puedan sugerir modificaciones en el contenido. De este modo, se hace posible ajustar al máximo los criterios de quienes crean los tesauros a las necesidades de quienes los utilizan.

El aumento del uso de Internet y la necesidad de sistematizar de manera coherente sus contenidos con vistas a facilitar la localización de los recursos que transporta, están dando un nuevo impulso a la aplicación de los sistemas tradicionales de clasificación de materias, que se emplean en las bibliotecas para crear directorios organizados por temas que asistan al usuario en la búsqueda y visualización de recursos electrónicos en la Web. La estructura jerárquica de estos sistemas permite una aproximación más sencilla a los temas que cubren. Si se han elaborado con funciones de hipertexto, es posible ir saltando de una parte a otra de la estructura de los contenidos. El usuario, al tener una idea más clara de los distintos grupos temáticos y sus subdivisiones, puede afinar las estrategias de búsqueda. Los sistemas jerárquicos tienen ahora un amplio campo de desarrollo en Internet gracias a la popularidad de la 'World Wide Web'. Cada vez son más numerosos los servicios que los utilizan. Pueden variar, desde los más universales, como la CDU o la LCC, hasta los esquemas de ámbito nacional o local, pasando por los especializados en una disciplina o área del conocimiento concreta y los expresamente diseñados para un servicio concreto de Internet, como 'Yahoo', 'Altavista' o 'Google'. Los tesauros en Internet, forma más elaborada de lenguaje documental que las clasificaciones, también pueden tener su utilidad práctica como instrumentos auxiliares de búsqueda en este tipo de recursos electrónicos en línea. Tienen las ventajas de las clasificaciones pero presentan sus términos de manera más completa y concreta. No obstante, esta particularidad los hace también algo más complicados de manejar que las simples clasificaciones jerárquicas.

#### 4. TESAUROS EN INTERNET

En Internet circulan ya numerosos tesauros en formato electrónico de libre acceso. Muchos de ellos permiten ser descargados por el propio usuario para utilizarlos en su ordenador fuera de línea. Otros, sin embargo, son tan extensos que hacen poco práctica su descarga o bien están protegidos de algún modo para evitarla. Unos son sencillos en su estructura y contenido, mientras otros son más complicados y requieren un cierto aprendizaje previo para su comprensión y manejo. Las áreas temáticas que cubren son también muy dispares, desde los aspectos más triviales hasta disciplinas científicas como la astronomía, la física, la informática o la medicina; además de ciencias sociales como la economía, la historia o el derecho. También pueden encontrarse clasificaciones y tesauros relacionados con la actividad de organizaciones e instituciones públicas o privadas y con la gestión de las Administraciones públicas. Algunos de estos tesauros permiten al usuario acceder a bases de datos a las que están asociados y en las que se referencian documentos sobre los temas de los que tratan los descriptores seleccionados, e incluso permiten recuperar los documentos en formato electrónico. Respecto a su autoría, gran parte de los tesauros de Internet corresponden a entidades públicas o privadas cuya actividad se relaciona con el tema del lenguaje documental que han desarrollado. Destacan las universidades y centros de investigación (CINDOC, NASA), así como las organizaciones internacionales (OCDE, OIT, UNESCO, Unión Europea).

Sobre la naturaleza y ventajas del uso de tesauros electrónicos se han publicado ya algunos trabajos, alguno de los cuales describe una muestra de tesauros seleccionados de Internet (A. de la Cueva Martín, 1999).

A continuación se describen a modo de ejemplo algunos de los más interesantes, tanto desde el punto de vista de su contenido como de su estructura, presentación y posibilidades de manejo.

##### **AAT (Art and Architecture Thesaurus)**

El vocabulario ATT (tesauro de arte y arquitectura), junto con sus auxiliares ULAN (lista unificada de nombres de artistas) y TGN (tesauro de nombres geográficos), constituye un lenguaje estructurado pensado para mejorar la gestión documental de información relativa al ámbito de las artes y la arquitectura. Se concibió como instrumento para ser utilizado como vocabulario controlado y de autoridades de las bases de datos sobre dichas materias que mantiene la fundación norteamericana J. Paul Getty Trust, institución responsable de su elaboración. Se compone de términos preferentes (descriptores) y sinónimos estructurados para catalogar, indizar o clasificar documentos.

El vocabulario AAT contiene cerca de 125.000 términos e informaciones complementarias en inglés con un alcance mundial y un espectro temporal que abarca desde la antigüedad hasta nuestros días. Cada uno de los términos preferentes del vocabulario se acompaña de un código numérico que lo identifica, mostrando también el conjunto de relaciones asociativas, jerárquicas o de sinonimia que mantienen con otros términos, además de notas aclaratorias sobre su definición y manejo. El ATT se estructura en una serie de grupos jerárquicos y facetas que proporcionan un esquema general de clasificación. Una faceta contiene una clase homogénea de conceptos dotados de unas características que los hacen distintos de los incluidos en otras facetas. Estas características diferenciales

se traducen en varias categorías: conceptos asociados, atributos físicos, estilos y períodos, agentes, actividades, materiales y objetos.

El ATT no sólo proporciona una completa terminología sobre disciplinas artísticas, sino que además los descriptores se acompañan de otros datos auxiliares que informan acerca de objetos, conceptos, artistas y lugares relativos a los conceptos que representan.

Se creó principalmente para un tipo de usuarios muy concreto: museos, bibliotecas especializadas, archivos, catalogadores de recursos visuales, realizadores de proyectos bibliográficos relacionados con las artes e investigadores sobre historia del arte. Su elaboración es el resultado de un proyecto de colaboración en el que han participado diversas instituciones.

Estamos ante un caso de lenguaje documental electrónico en línea caracterizado por su especialización y su gran extensión, de modo que en realidad no se trata de un vocabulario único sino de tres que se complementan para un mismo propósito general: la descripción temática de documentos relacionados con el arte. Su creador es una institución privada, la fundación J. Paul Getty Trust, que concede licencias de explotación de los vocabularios en diversos formatos a aquellas entidades culturales que lo soliciten. Por lo tanto están protegidos por derechos de 'copyright'. Debido a sus grandes dimensiones, estos vocabularios presentan una estructura más compleja que la mayoría de los tesauros. No obstante, la navegación a través de su estructura está asistida en cada vocabulario por motores de búsqueda que dan acceso directo a los términos consultados. Otro valor añadido radica en el apoyo que presta a los usuarios, proporcionándoles una información muy completa sobre las características y modo de utilización de los tesauros a través de su página Web <<http://www.getty.edu/research/tools/vocabulary/aat/>>.

### **Tesoro de Educación**

Creado por el Departamento de Análisis y Planificación de la Universidad Complutense de Madrid, este tesoro monolingüe (español) especializado en el campo de la educación, tiene la particularidad de mostrar una interfaz en forma de motor de búsqueda, permitiendo realizar búsquedas por términos exactos o derivados y utilizar operadores 'booleanos' y truncamientos para localizar descriptores.

En cuanto a su presentación, hay que destacar que no ofrece nada más que el motor de búsqueda y una muy escueta ayuda. No tiene parte introductoria y por tanto no ofrece detalles sobre aspectos como su proceso de elaboración, estructuración, temática o instrucciones de uso. Tampoco muestra la estructura jerárquica ni la alfabética, sin embargo, al acceder a cada descriptor se presentan las relaciones jerárquicas o asociativas que mantienen con otros términos del tesoro, permitiendo además la posibilidad hipertexto de acceder directamente a ellos con un 'clic'.

Por otro lado, si bien es cierto que el motor de búsqueda tiene la ventaja de recuperar rápidamente los términos requeridos por el usuario, el hecho de que éste no pueda visualizar la estructura de todo el tesoro para desplazarse por ella le impide hacerse una idea general del mismo y limita la utilidad práctica del lenguaje documental, al no facilitar un conocimiento integral de su contenido que permita un mejor aprovechamiento de las posibilidades del tesoro. Si el usuario pudiera conocer la estructura global del tesoro y saber de qué manera están dispuestos los términos, podría hacer un uso más eficaz del mismo.

### **Tesoro Europeo de la Educación (TEE)**

Se trata de un tesoro especializado en educación cuya versión electrónica disponible actualmente en Internet data de 1998, siendo sus creadores la Comisión Europea y el Consejo de Europa. En su elaboración han participado diversos especialistas del ámbito educativo y su adaptación a las necesidades de los usuarios es permanente. Para desarrollarlo se han tenido en cuenta los criterios de la norma ISO 5964-1985 sobre 'Principios rectores para la elaboración y el desarrollo de tesauros multilingües'. Está publicado en 11 lenguas oficiales de la Unión Europea, cada una de las cuales es tratada con mismo nivel de relevancia, circunstancia que lo convierte en un importante instrumento de intercambio de información sobre temas educativos entre diversos países.

Se utiliza para la indización de la base de datos del Consejo de Europa 'EUDISED' (European Documentation and Information System in Education) y para la base de datos 'EURYDICE' de la Unión Europea sobre aspectos relacionados con la educación. Comprende las siguientes materias: principios y sistemas educativos, políticas de enseñanza, establecimientos de enseñanza, personal docente y alumnos; edificios escolares y materiales pedagógicos, programas de enseñanza y materias impartidas, investigación en materia de educación e información pedagógica, psicología de la educación, sociología de la educación, economía de la educación, y administración pública. Asimismo, se recoge una lista de nombres de países y regiones.

La estructura del tesoro es la típica de los tesauros tradicionales en formato papel: introducción, parte sistemática y parte alfabética; con descriptores, no-descriptores que remiten a los términos admitidos (sinónimos o cuasisinónimos), notas de alcance, relaciones jerárquicas, asociativas y de pertenencia a un microtesoro, así como de equivalencia lingüística. La edición accesible desde Internet se localiza en la página Web de EURYDICE <<http://www.eurydice.org/TeeForm/>> y con ella se pretende que sea accesible al mayor número posible de usuarios. Estos tienen además la posibilidad de expresar sus opiniones y sugerencias de manera interactiva a través de la propia página Web. Se presenta en varios archivos de formato PDF y sin estructura de hipertexto, con las limitaciones en cuanto a manejabilidad que ello supone, y está permitida la descarga parcial o total de su contenido. En definitiva, se trata de un tesoro electrónico pero con el aspecto de un tesoro en formato papel.

### **Tesoro EUROVOC**

El tesoro electrónico 'EUROVOC' <<http://europa.eu.int/celex/eurovoc/index.htm>>, publicado también en formato impreso y en disco óptico, es un lenguaje documental multilingüe redactado en 15 idiomas oficiales de países miembros de la Unión Europea que recopila términos referidos a la actividad de las instituciones comunitarias. Ha sido desarrollado de manera conjunta por el Parlamento Europeo, un comité de la Comunidad Europea y la Oficina de Publicaciones Oficiales de las Comunidades Europeas, contando con el apoyo de la Dirección General XIII. Su mantenimiento corre a cargo de la Oficina de Publicaciones Oficiales. Hasta el momento se han realizado cuatro ediciones de EUROVOC, la última de las cuales es la que se halla disponible en Internet en formato PDF. Se concibió como instrumento de indización y búsqueda para bibliotecas y centros de documentación e información especializados en derecho, legislación y política. La Unión Europea lo utiliza en la biblioteca del Parlamento Europeo, en la propia Oficina de Publicaciones Oficiales y en otras instituciones comunitarias. Su uso se extiende también



a centros de documentación de diversos parlamentos nacionales y terceros países. En la elaboración del tesoro se han seguido estrictamente las pautas establecidas por la Organización Internacional de Normalización para la creación de tesoros monolingües (ISO 2788-1986) y multilingües (ISO 5964-1985).

El contenido temático se distribuye en 21 áreas que contienen descriptores sobre: vida política, relaciones internacionales, Comunidades Europeas, derecho, intercambios comerciales, asuntos financieros, asuntos sociales, educación y comunicación, ciencia, empresa y competencia, trabajo y empleo, transportes, medio ambiente, agricultura, silvicultura y pesca, sector agroalimentario, producción tecnología e investigación, energía, industria, términos geográficos y organizaciones internacionales. Cada uno de estos grandes grupos se divide en varios subgrupos llamados 'microtesoros' (127 en total). El tesoro se divide en tres partes fundamentales, repartidas en diversos archivos PDF: parte introductoria, parte sistemática y parte alfabética permutada; y cuenta con más de 6.500 descriptores.

Se trata de un tesoro muy completo y útil para el control de información general sobre diversos aspectos del ámbito socioeconómico y otros campos como el derecho, la educación o el medio ambiente. Su estructura responde al mismo modelo de los tesoros clásicos editados en papel, de hecho el aspecto visual que ofrece es idéntico al de la versión impresa. Los descriptores presentan todas las relaciones semánticas propias de un tesoro tradicional: de pertenencia a un microtesoro, de equivalencia, jerárquicas y asociativas, incluyendo notas de aplicación con aclaraciones sobre la definición y uso de algunos términos. Cada descriptor tiene asociado un código numérico que indica el microtesoro al que pertenece para facilitar la clasificación documental.

El tesoro EUROVOC es de libre acceso y está permitida la descarga de todo su contenido, lo cual representa una gran ventaja para su posible aprovechamiento por cualquier usuario o centro de documentación interesado en utilizarlo o adaptarlo con fines propios. Sin embargo, en cuanto a manejabilidad, sus prestaciones no incluyen funciones hipertexto debido a las limitaciones propias del formato PDF, de manera que las búsquedas de términos dentro del tesoro se deben realizar por localización secuencial. En esto se parece a los tesoros impresos.

En definitiva, EUROVOC es un tesoro muy extenso, completo y normalizado, pero su manejo es poco práctico debido a que su presentación en formato PDF no incluye funciones de control hipertexto para facilitar la movilidad a través de su estructura.

### **Tesoro IRIS ('IRIS Keyword Thesaurus')**

El tesoro IRIS ha sido desarrollado en lengua inglesa por la Universidad de Illinois Urbana-Champaign y contiene cerca de 2.100 descriptores que sirven como puntos de acceso para lanzar consultas directas a la base de datos bibliográfica IRIS a través de un motor de búsqueda. Los términos abarcan campos muy amplios y diversos repartidos en 45 grupos temáticos que comprenden áreas como agricultura, arte, ciencias del comportamiento, biotecnología, administración de empresas, comunicación, informática, educación, energía, medio ambiente, derecho, ciencias médicas, lugares geográficos, matemáticas, humanidades, ciencia, transportes, urbanismo, etc.

Su estructura consta de un índice sistemático al comienzo, donde aparecen los grandes grupos de temas, y de una parte alfabética en la que se ordenan todos los descriptores. No dispone de una auténtica parte introductoria en la que se describan detalles acerca de su

elaboración e indicaciones de uso. El tesoro tiene formato de documento HTML, ordenando los términos de manera alfabética y desplegando toda la jerarquía de términos que contienen.

Las relaciones que refleja son las de sinonimia (USE, See under) y asociación (RT). Los términos se acompañan también de un número secuencial a su izquierda que facilita la realización de consultas abreviadas a la citada base de datos. Se trata de un tesoro estructurado de manera muy sencilla y con escasos vínculos internos para enlazar términos entre sí o moverse rápidamente a través del documento. Esta sencillez hace que su uso sea bastante simple y eficaz, aunque se echa en falta una estructura sistemática que facilite ver todos los descriptores contenidos en cada grupo temático.

Es un buen ejemplo de lenguaje documental sencillo y práctico, aunque no aplica todas las posibilidades de la norma ISO para la construcción de tesauros y adopta una estructura de hipertexto poco desarrollada. Por ello, habría que considerarlo más como un sistema de clasificación que como un verdadero tesoro. Se localiza en la siguiente dirección: <<http://www.carousel.lis.uiuc.edu/~iris/thesaurus.html>>.

### **Tesoro UNBIS**

Este tesoro, que en formato electrónico supone la cuarta edición desde su aparición en 1981, contiene la terminología utilizada para el análisis documental del Sistema de Información Bibliográfica de las Naciones Unidas (UNBIS), siendo adoptado también como léxico documental por el Sistema Oficial de Documentos de la ONU.

Ha sido desarrollado en los cinco idiomas oficiales de la ONU (árabe, chino, inglés, francés, ruso y español). Es un buen ejemplo de tesoro multilingüe que ha superado con éxito las dificultades técnicas que plantea elaborar un lenguaje documental en distintos alfabetos. Su cobertura temática comprende una amplia variedad de términos relacionados con la actividad desarrollada por dicha organización internacional: política, economía, sociedad, agricultura, industria, transportes, comercio, demografía, salud, empleo, cultura, ciencia y tecnología, etc. Los descriptores son seleccionados siguiendo dos criterios fundamentales: la consistencia con la terminología utilizada por la ONU en sus documentos, y la compatibilidad con los términos contenidos en otros tesauros utilizados en el Sistema de Naciones Unidas para facilitar el intercambio de información entre las distintas organizaciones que lo integran (Banco Mundial, OIT, OMS, FAO, OIEA, Unesco).

El Tesoro UNBIS es ampliado y actualizado continuamente. Se presenta formando parte de la página Web del Sistema de Información Bibliográfica de las Naciones Unidas <<http://www.unhq-appspub-01.un.org/LIB/DHLUNBISThesaurus.nsf/>>. Tiene el aspecto de documento HTML.

Las opciones de consulta de términos son muy completas gracias al uso de motores de búsqueda. Un primer motor de búsqueda, permite recuperar términos tecleando directamente su nombre, admitiéndose también el uso de truncamientos, o bien realizar búsquedas lógicas mediante operadores 'booleanos' o con frases completas. Un segundo motor de búsqueda permite la búsqueda alfabética de términos. Asimismo, se pueden desplegar las distintas categorías temáticas en las que se agrupan los descriptores dentro de la estructura jerárquica del tesoro.

Dichas opciones se activan mediante una serie de enlaces situados en el marco superior de la página Web, donde también se incluye un enlace de ayuda sobre el manejo del tesoro y otro que ofrece información adicional sobre las características del tesoro. Otros

vínculos enlazan el tesoro con el propio sistema de información bibliográfica para el que fue diseñado y del que forma parte, así como con las versiones en los demás idiomas oficiales de Naciones Unidas. Un aspecto interesante es el enlace que muestra un listado de otros tesauros y terminologías afines a las que es posible acceder directamente activando un hipervínculo. Además, el usuario puede enviar sugerencias por correo electrónico a los responsables del mantenimiento del tesoro.

Las relaciones entre los descriptores son de tipo asociativo, jerárquico y de equivalencia. Contiene notas de alcance y, cuando se recupera un término en modo de búsqueda alfabética, se puede desplegar la estructura completa de las relaciones que tiene asignadas en todos los idiomas del tesoro.

Cada descriptor tiene asignado un código numérico, compuesto por tres pares de dígitos separados por puntos, que corresponde a la categoría temática a la que pertenece dicho término y tiene utilidad a efectos de clasificación.

Consideramos que el Tesoro UNBIS ofrece un diseño claro y sencillo que facilita su utilización por cualquier tipo de usuario. Las posibilidades propias de un documento hipertexto son bien aprovechadas en este caso, de modo que favorece la navegación a través de su contenido textual.

### **Macrothesaurus de la OCDE**

La versión electrónica de este tesoro de términos económicos disponible en Internet se presenta en formato HTML con estructura hipertexto y en cuatro idiomas: inglés, francés, español y alemán. La Organización para la Cooperación y el Desarrollo Económicos (OCDE) es la institución responsable de su elaboración, publicación y mantenimiento.

Su acceso es público y, tanto por la cantidad de términos que recoge como por la diversidad de áreas y conceptos que cubre, se trata de uno de los mejores tesauros especializados en economía que pueden encontrarse en la Red. Consta de una parte sistemática en la que figuran los 19 grandes grupos temáticos en los que se divide. Cada uno de ellos es un hiperenlace que al activarse abre la jerarquía de términos que contiene. Estos, a su vez, también son hiperenlaces que al activarse remiten a la descripción del término (con todas sus relaciones y notas explicativas) contenida en la parte alfabética.

A esta parte del tesoro se puede acceder también directamente y en ella aparecen todos los términos por orden alfabético y conteniendo todos los detalles que se refieren a su ubicación en la estructura del tesoro y las relaciones con otros términos (grupo temático al que pertenecen, relaciones jerárquicas, asociativas o de sinonimia, notas de alcance, código identificador). Las consultas al tesoro pueden realizarse navegando a través de su contenido o bien mediante el uso de un motor de búsqueda. Por sus características, es un tesoro bastante completo, sencillo y de fácil manejo, siendo de gran utilidad para usuarios dedicados al ámbito de la economía. Este sería un buen ejemplo de tesoro típico especializado en un área concreta del conocimiento. Este tesoro puede localizarse tecleando la siguiente dirección: <<http://info.uibk.ac.at/info/oecd-macroth/>>.

### **Tesoro OIT <<http://www.ilo.org>>**

El Tesoro OIT ha sido elaborado a instancias de la Organización Internacional del Trabajo, perteneciente al sistema de Naciones Unidas, como instrumento de indización y consulta del catálogo automatizado de su biblioteca central y otros servicios de información. Recopila más de 4.000 términos relativos al ámbito laboral y al desarrollo económi-

co y social, todos ellos expresados en cuatro lenguas: inglés, francés, alemán y español; y repartidos en 19 facetas temáticas que se dividen de manera jerárquica en diversos grupos y abarcan aspectos como la política de empleo y de trabajo, la planificación de recursos humanos, las normas de trabajo y relaciones de trabajo, la formación profesional, el desarrollo económico y social, la seguridad social, las condiciones de trabajo, salario, salud y seguridad en el trabajo, la promoción empresarial y el desarrollo gerencial. El Tesauro OIT surge de la evolución de una lista alfabética de más de 1.200 palabras claves creada en 1964.

Se presenta en dos partes principales: una primera consistente en cuatro índices alfabéticos, uno por cada idioma representado; y una segunda sistemática en la que se despliega la totalidad de los descriptores a lo largo de las diversas facetas. A los descriptores se les asigna una numeración secuencial para facilitar el reenvío de la parte alfabética a la parte sistemática e indicar el tema principal del documento indizado de forma codificada. Este código identifica la clasificación temática con el fin de facilitar la recuperación automatizada de los documentos indizados y aumentar las posibilidades de precisión de las búsquedas combinándolos con los descriptores. De este modo, el tesauro puede ser usado también como una clasificación, no sólo como un lenguaje documental postcoordinado.

Las conexiones entre sus términos son señaladas mediante relaciones de sinonimia, asociación, jerarquía y equivalencia. Los descriptores pueden incluir también notas de alcance cuando se considera oportuno precisar su definición o dar indicaciones sobre su uso. Asimismo, se señalan los términos recientemente añadidos al tesauro por medio de un asterisco y el año de su incorporación.

Se actualiza periódicamente, añadiendo nuevos términos y suprimiendo otros desfasados o en desuso.

### **Clasificación de la Biblioteca del Congreso ('Library of Congress Classification outline')**

Clasificación de materias de la Biblioteca del Congreso de Washington publicada en Internet en inglés. Por la temática de sus términos, tiene un ámbito generalista, abarcando áreas como la filosofía, psicología, religión, historia, geografía, ciencias sociales, política, leyes, educación, música, arte, lenguaje y literatura, ciencia, medicina, agricultura, tecnología, ciencias militares y navales, y ciencias de la información.

No se trata de un tesauro propiamente dicho sino de una clasificación de materias accesible en formatos HTML, PDF y Word. Los términos se presentan en una estructura jerárquica organizada en 21 grandes capítulos que van desglosándose en apartados y subapartados. Esta estructura es navegable en su totalidad mediante enlaces hipertexto y los términos tienen asociados también códigos alfanuméricos para su representación abreviada. La página Web de la Biblioteca del Congreso es de gran interés para los documentalistas por la gran cantidad de información que ofrece sobre catalogación. Es este un ejemplo de lenguaje documental electrónico en Internet de gran extensión y carácter enciclopédico, propio de una biblioteca nacional. Por eso es más práctico su manejo en línea, ya que descargarlo en un ordenador local requiere tiempo y bastante espacio en memoria y se organiza en varios archivos. Se aloja en la siguiente dirección: <<http://www.lcWeb.loc.gov/catdir/cpsol/lcco/lccowp.html>>.

## Tesoro de Legislación Histórica de España

El 'Tesoro de Legislación Histórica de España' es un claro ejemplo de lenguaje documental que adopta la forma de una clasificación jerarquizada de términos muy sencilla, a pesar de su denominación de tesoro.

Se presenta en formato PDF y sin ningún tipo de enlaces internos que faciliten la movilidad a través de su estructura. Fue desarrollado para asistir en sus búsquedas a los usuarios de la base de datos en Internet que recopila los principales textos de la legislación histórica de España desde los códigos góticos de la Edad Media hasta las Cortes de Cádiz. La base de datos es el resultado de un proyecto de colaboración entre la Universidad Autónoma de Madrid y el Ministerio de Cultura, en cuyo sitio Web se encuentra alojada <<http://www.cultura.mecd.es/archivos/lhe>>. El tesoro conduce a la descripción de los documentos incluidos en la base de datos y a sus imágenes.

Se trata de una clasificación muy básica pero que puede resultar interesante para quienes deseen elaborar un sistema de clasificación sencillo pero eficaz. Por ello, es recomendable como modelo para la realización de sistemas de clasificación sobre temas muy concretos que no requieran sistematizar una gran cantidad de términos.

## Clasificación JEL

La revista *Journal of Economic Literature* (JEL), publicada por la American Economic Association (AEA), es una de las publicaciones periódicas especializadas en Economía más prestigiosas del mundo. Para indizar y recuperar sus artículos ha desarrollado un sistema de clasificación jerárquica a tres niveles que ha alcanzado amplia difusión internacional, de manera que numerosas revistas del mismo ámbito temático la utilizan también como referencia acompañando a las entradas o al resumen de los artículos que publican.

Los términos de esta clasificación se dividen en veinte grandes áreas referidas a distintas materias de la Economía, cada una de las cuales se divide en grupos secundarios. A su vez, estos últimos contienen una serie de términos configurando un tercer nivel jerárquico. Cada término puede expresar uno o varios conceptos y se encuentra representado por un código alfanumérico que consta de una letra mayúscula inicial, indicativa del grupo de primer nivel al que pertenece, y de un número de uno o dos dígitos que señala la pertenencia a alguno de los otros dos niveles jerárquicos inferiores. Son estos códigos los que se utilizan para indizar y permitir la búsqueda de los artículos y documentos accesibles a través de la página web que la propia revista tiene en Internet. Los códigos son asimismo utilizados por las publicaciones económicas que adoptan la Clasificación JEL para asignar conceptos claves a sus artículos.

Es una clasificación muy sencilla que es posible localizar también en otras páginas y sitios web. En España, algunas de las publicaciones económicas más destacadas la utilizan, como *Papeles de Economía Española*, *Boletín Económico ICE*, *Revista de Economía ICE*, *Investigaciones Económicas y Hacienda Pública Española*, entre otras. Aunque circulan traducciones en otros idiomas, como en español, la versión original de la Clasificación JEL es la inglesa y actualmente se puede encontrar en formato HTML en la dirección de Internet de la propia AEA <[http://www.aeaweb.org/journal/jel\\_class\\_system.html](http://www.aeaweb.org/journal/jel_class_system.html)>.

### **Tesaurus ISOC de Economía**

El productor de este tesaurus es el CINDOC, entidad española de carácter público integrada en el Consejo Superior de Investigaciones Científicas (CSIC). Se trata de un lenguaje documental diseñado para permitir una indización homogénea de los documentos incorporados a la base de datos ECOSOC del CSIC. Recoge más de 6.600 términos, de los cuales unos 5.300 son descriptores y cerca de 1.300 no descriptores. Incluye además 240 notas de alcance. Los descriptores muestran también su traducción en inglés y francés.

Se trata de un vocabulario controlado muy útil para centros de documentación especializados en Economía y que surgió como herramienta en formato impreso, pasando posteriormente a formato electrónico en línea (HTML) con la expansión de Internet. La página Web a la que se accede para su consulta muestra una pantalla dividida en dos franjas verticales, a la izquierda del usuario aparece un índice con tres apartados cuyo contenido es mostrado en la franja derecha al activarse un hipervínculo: versión impresa (muestra la ficha catalográfica y el Depósito Legal del tesaurus impreso), versión electrónica (muestra una introducción explicativa sobre el tesaurus) y búsquedas (abre otros dos apartados: el índice KWOC y el índice jerárquico del tesaurus). En la parte superior izquierda de la pantalla aparecen cuatro hipervínculos que dan acceso al índice, a la página de inicio, a un motor de búsqueda por términos y al índice KWOC. Se divide en 13 áreas temáticas. De la 01 a la 10 son específicas de Economía y se basan en la clasificación del 'Journal of Economic Literature' (JEL). La 11 y la 12 agrupan disciplinas relacionadas con la Economía y la 13 se ocupa de términos sobre Documentación. Para su elaboración se han seguido las 'Directrices para el Establecimiento y Desarrollo de Tesaurus Monolingües' de las normas PGI/81/WS/15 y UNE 50-106-90.

El tesaurus está constituido por descriptores y no descriptores que presentan un sistema de relaciones de equivalencia, jerárquicas y asociativas. Se han excluido del mismo los nombres o identificadores (nombres propios, instituciones, períodos históricos, monedas nacionales, etc.), que figuran en un anexo aparte. Las partes que integran el tesaurus ISOC son: introducción, parte alfabética, parte sistemática e índice permutado KWOC, además del citado anexo de identificadores. El manejo de este completo vocabulario especializado es sencillo y permite un acceso rápido al término consultado gracias a una estructura hipertexto bien definida. No obstante, creemos que estos aspectos podrían mejorarse para hacerlo todavía más amigable y accesible.

Este tesaurus constituye un buen ejemplo de lenguaje documental electrónico en línea creado en España por una institución pública investigadora en formato HTML y dedicado a un área especializada del conocimiento y sin restricciones de acceso, aunque no es factible su descarga en modo local. Además de su distribución comercial en papel, también existe una versión publicada en CD-ROM.

Para profundizar en el conocimiento del proceso de desarrollo de este y de otros tesaurus electrónicos para la Web del CINDOC, resulta de utilidad consultar el artículo publicado por A. Valle Bracero, et al. (2000).

### **TESAMAT (Tesaurus Ciencias Matemáticas)**

Este tesaurus especializado en matemáticas es el resultado de varios años de colaboración de diversos expertos universitarios. Su publicación corre a cargo de la Facultad de Ciencias Matemáticas de la Universidad Complutense de Madrid y está disponible en la página web de la Biblioteca de la UCM <<http://www.ucm.es/BUCM/documentosBUC>>.

La estructura de TESAMAT consta de cinco partes. La primera es una introducción en la se que describen brevemente su proceso de elaboración y contenido. En segundo lugar, una relación terminológica relativa al campo de las matemáticas extraída de los esquemas básicos de tres clasificaciones de materias de ámbito internacional: la CDU (Clasificación Decimal Universal), la AMS (Mathematical Subject Clasification of the American Mathematical Society) y la LC (Clasificación of the Library of congreso). La parte tercera incluye el listado de los descriptores que componen el tesoro propiamente dicho ordenados alfabéticamente. Los términos admitidos, denominados en este caso encabezamientos de materias, se presentan con todas las relaciones posibles de los tesauros al uso: traducción inglesa del término válido como descriptor, notas de alcance, relaciones de sinonimia y equivalencia, términos genéricos, términos específicos, términos relacionados y correspondencias con los códigos de los términos afines de las clasificaciones CDU, AMS y LC. La cuarta parte contiene un glosario inglés-español y la quinta y última las equivalencias de los códigos correspondientes a los términos de las citadas tres clasificaciones consideradas.

El tesoro TESAMAT editado en formato electrónico se haya contenido en un único archivo PDF y no presenta elementos de navegación interna que permitan ciertas funciones hipertexto. Por tanto, su consulta sólo es posible de manera secuencial, tal como se hace con un libro impreso. En este sentido, se trata de un documento muy limitado, ya que su manejo no es ágil y no admite que los términos sean adaptados a sistemas locales de información. Parece pensado más para su uso como documento impreso que para ser utilizado como recurso electrónico.

Por su parte, tanto las referencias a concordancias con otras clasificaciones como la inclusión del glosario, constituyen una novedad respecto a la estructura del resto de los tesauros considerados en este artículo.

## CONCLUSIONES

El formato electrónico se presta muy bien a la elaboración y presentación de lenguajes documentales para su uso en sistemas informatizados, bien en modo local o en red. Permite sustituir publicaciones de gran volumen en papel por un soporte que facilita una consulta más ágil. La accesibilidad y la manejabilidad se ven notablemente incrementadas si se añaden propiedades hipertexto al tesoro, clasificación o lista de materias.

Algunos tesauros y vocabularios electrónicos sólo se encuentran en Internet. Otros suelen publicarse además en soporte impreso o en CD-ROM, como es el caso del tesoro EUROVOC.

Los tesauros en Internet pueden aparecer como documentos individuales que no dan acceso directo a ningún sistema documental, o bien formado parte de un sistema documental en línea como instrumento auxiliar que permite realizar búsquedas de referencias en sus bases de datos e incluso localizar documentos electrónicos. Como ejemplo de ello tenemos el Tesoro UNBIS.

Muchos tesauros se pueden bajar de Internet y copiarlos en nuestro disco duro o en otro soporte de almacenamiento masivo. De este modo, es posible utilizarlos fuera de línea. Sin embargo, aquellos que sólo son accesibles en línea (debido normalmente a su extensión o a que están protegidos por 'copyright' para evitar copias no permitidas), no pueden ser descargados de la Red a otro ordenador. Su consulta supone un gasto en tiem-

po de conexión cada vez que queramos acceder al sitio Web donde se aloja. En estos casos hay que asumir el riesgo de que no siempre puede encontrarse disponible el recurso, bien por saturación de la Red en ese momento o por problemas técnicos del servidor, de nuestro proveedor Internet, de nuestra red corporativa o de nuestro propio equipo informático.

Frente a los tesauros en papel, los electrónicos requieren por parte del usuario un cierto conocimiento previo del manejo de un ordenador y del uso de Internet.

Existen condicionamientos técnicos de los que depende un correcto uso de los tesauros electrónicos, sean en línea o fuera de línea: disponer de una infraestructura informática adecuada, tanto en lo referente a los equipos y soportes como a los programas. La obsolescencia tecnológica es muy rápida en la industria de la informática y las telecomunicaciones. Además, los formatos en los que se registran los documentos electrónicos también sufren cambios y renovaciones, pudiendo incluir elementos como motores de búsqueda o pequeñas aplicaciones que desarrollan acciones adicionales.

El formato electrónico es ideal para presentar un lenguaje documental en Internet con las características de un documento hipertexto. Varias son sus ventajas: sencillez y rapidez de manejo, actualización automática, difusión telemática, bajo coste de edición en comparación con el formato papel, consulta simultánea por más de un usuario, posibilidades de adaptación y modificación del texto en modo local para elaborar un lenguaje documental propio ajustado a necesidades informativas concretas, etc.

Se ha visto también como algunos tesauros pueden ser utilizados indistintamente como clasificaciones o como tesauros propiamente dichos, según se usen sus términos en la forma codificada o literal.

En general, cuanto mayor es la institución que elabora o patrocina la edición de un tesoro, más amplio suele ser su contenido y más complicada su estructura, incluyendo muchas veces motores de búsqueda que facilitan la localización directa de términos y descriptores. Este sería el caso del Macrothesaurus de la OCDE o del tesoro AAT. La amplitud y la profundidad de estos lenguajes documentales los hacen poco prácticos para su manejo en modo local debido al gran espacio en memoria que requieren y a la cantidad de términos que contienen. Además, algunos están protegidos y no pueden copiarse o precisan de una licencia para su explotación. También pueden presentar incompatibilidades con los sistemas locales donde se pretenden instalar.

Un tesoro electrónico de estructura compleja colgado en Internet y cuyo acceso se realiza a través de multitud de pantallas puede producir en el usuario una fatigosa saturación de imágenes, haciéndolo incomodo y poco práctico. Hay que tener en cuenta que se está mirando a un objeto luminoso como es una pantalla de ordenador y por tanto es más molesto que fijar la atención sobre un documento de papel.

Los lenguajes documentales en formato electrónico se muestran muy útiles cuando se trata de terminologías en varios idiomas reunidas en un único tesoro o clasificación, puesto que gracias a las propiedades hipertexto se puede saltar de un término a su equivalente traducido en otro idioma. Claros ejemplos de ello son el Tesoro EUROVOC y, sobre todo, el Tesoro UNBIS, ambos asociados a la actividad documental de grandes organizaciones internacionales.



**BIBLIOGRAFÍA**

- Codina, Lluís. Evaluación de recursos digitales en línea: conceptos, indicadores y métodos. *Revista Española de Documentación Científica*. 2000, vol. 23, nº 1, p. 9-44.
- Cueva Martín, Alejandro de la. Acceso y utilización de tesauros en Internet. *Revista Española de Documentación Científica*. 1999, vol. 22, nº 4, p. 531-540.
- García Camarero, Ernesto; García Melero, Luis Angel. La biblioteca digital. Madrid: Arco/Libros, 2001, p. 381.
- Martín González, Juan Carlos; Merlo Vega, José Antonio. Las revistas electrónicas: características, fuentes de información y medios de acceso. *Anales de Documentación*. 2003, nº 6, p. 155-186.
- Russell, Rosemary; Day, Michael. *HILT: High-Level Thesaurus project* [en línea]: *automated and manual approaches to the provision of thesauri and subject vocabularies*. Ver. 1.0. 5 october 2001. Bath, UK: University of Bath. <<http://www.ukoln.ac.uk/hilt/interfaces/>>.
- Valle Bracero, Antonio; Rey Guerrero, Alfredo del; Páez Mañá, Jorge y Valle Bracero, Reyes. Tesauros en HTML: un modelo de diseño y estructura para su consulta en la malla mundial (WWW). *Revista Española de Documentación Científica*. 2000, vol. 3, nº 2, p. 159-178.