



Temas de Biblioteconomía

Lenguajes documentales
Principales tipos de clasificación
Encabezamientos de materia,
descriptorios y tesauros

Autor: César Martín Gavilán

Fecha: 11/04/09

Lenguajes documentales

La producción documental ha aumentado exponencialmente en los últimos años, algo que ha generado nuevas necesidades, nuevos sistemas y técnicas de tratamiento documental. Los lenguajes documentales surgen de la necesidad de conseguir que el emisor y el receptor de la información lleguen a el entendimiento en un contexto en el cual cada vez hay más material y, por lo tanto, más ruido que dificultan la comunicación. De ahí la necesidad de crear códigos como elementos secundarios, nuevos productos a partir de la obra original. Esa respuesta anticipada a un riesgo potencial de falta de comunicación constituye de hecho la esencia de todo el esfuerzo del análisis documental.

El proceso se podría representar de la siguiente manera: el autor crea la obra con un código determinado, el analista (o agentes software de información) lo traduce a su código (adaptado para la búsqueda y recuperación de la información) y el receptor utiliza ese nuevo código para encontrar la obra original.

Los lenguajes documentales son, pues, los que establecen la comunicación entre el documento original y el usuario, haciendo de puente o intermediario entre las informaciones contenidas en los documentos y las informaciones solicitadas por los usuarios.

Definición de lenguaje documental

C. Guinchat y M. Menou (1984) definen los lenguajes documentales como "lenguajes convencionales utilizados por una unidad de información para describir el contenido de los documentos para almacenarlos y recuperar las informaciones".

Blanca Gil Urdiciáin (1996) entiende por lenguaje documental "todo sistema artificial de signos normalizados, que facilitan la representación formalizada del contenido de los documentos para permitir la recuperación, manual o automática, de información solicitada por los usuarios".

Quizás la definición más ajustada de lenguaje documental sería la de sistema artificial de signos utilizados en las operaciones de indización que permite la representación de contenido documental para su posterior recuperación, sirviendo como medio para la interrogación, recuperación y difusión de información pertinente para el usuario. Al fin y al cabo, la finalidad última de los lenguajes documentales es la generación del conocimiento en el usuario, algo que sólo será posible si éste consigue llegar al documento primario.

El lenguaje documental es considerado objeto de estudio de la lingüística documental, cuyas bases teóricas proceden de disciplinas como la lingüística,

la semántica, la sintaxis, la onomasiología, la lexicología y la lingüística textual, entre otras, y está relacionada con otras áreas de conocimiento como la inteligencia artificial, la lógica e incluso con la teoría de la comunicación y de la información.

El lenguaje documental no es un “lenguaje natural”, aunque se sirva de signos del lenguaje natural (palabras) que representan nuestro conocimiento de la realidad. Los lenguajes naturales presentan problemas de difícil solución, como la polisemia o la sinonimia, que los hacen poco adecuados para la indización, lo cual no quiere decir que el lenguaje natural no se utilice en algunos tipos de indización. Los lenguajes documentales surgen precisamente para reducir los problemas que produce la indización con lenguajes naturales. De esta forma, los términos “lenguaje de indización” y “lenguaje documental” no son sinónimos: todos los lenguajes documentales son lenguajes de indización, pero no todos los lenguajes de indización son lenguajes de documentales. Las listas de palabras clave y descriptores libres no son lenguajes documentales.

En el lenguaje documental, las palabras adquieren valor semántico por medio de su normalización y de las reglas morfosintácticas que lo articulan. Son precisamente esas reglas las que le dan categoría de “lenguaje artificial”. Además, en ocasiones el lenguaje documental reemplaza esos signos del lenguaje natural por símbolos cargados de significado preciso, de forma que lleva a cabo la representación convencional del documento a partir de una correspondencia analógica; éste es el caso de los sistemas de clasificación que, por el procedimiento de notación, representan ideas de manera unívoca a través de, por ejemplo, números (CDU, Dewey ...).

Es justo reconocer los límites e inconvenientes del lenguaje controlado frente al lenguaje libre: los lenguajes controlados permiten una indización menos específica y exhaustiva que el lenguaje libre, su actualización es difícil (necesitan un estudio y análisis previo a la incorporación de nuevos términos), encarecen los procesos de análisis documental y abren la posibilidad a errores puesto que la elección de puntos de acceso está en manos de profesionales y no de los autores. Sin embargo, todas las ventajas que incorporan los lenguajes libres en la fase inicial e intermedia de la cadena documental, se invierten al final, dificultando la recuperación de la información puesto que no se controlan los problemas de polisemia, homonimia y sinonimia del lenguaje natural, así como tampoco se resuelve el problema del multilingüismo. Quizás la mejor solución es la combinación de ambos lenguajes en las distintas fases de la cadena documental.

La renuncia a los lenguajes documentales trae sin lugar a dudas limitaciones evidentes. ¿Cómo se evalúa la cobertura temática de una colección de revistas científicas de una universidad si no existen encabezamientos de materia ni se ha utilizado clasificación alguna? Pagando el servicio de Ulrich's Serials

Analysis System, por ejemplo, y confiando en el uso preciso del ISSN en los registros bibliográficos.

Tipos de lenguajes documentales

Existen diversos criterios de clasificación de los lenguajes documentales, aunque los más generalizados son aquellos que utilizan como criterio el control ejercido sobre el vocabulario, la coordinación de los términos (en qué momento se combinan los elementos) y la estructura.

Dependiendo del grado de control terminológico, se pueden distinguir:

- *Lenguajes libres*. Lenguajes fundados en el principio de postcoordinación, se componen de términos no predefinidos que se va generando a partir de la realización de procesos de indización. Lenguajes altamente pertinentes, coherentes, de bajo coste y que proporcionan una enorme cantidad de resultados cuando se realizan búsquedas. De este tipo son las listas de descriptores libres y las listas de palabras clave (Uniterm, Índices permutados, Precis ...). Los lenguajes libres no son propiamente lenguajes documentales puesto que para que reciban este nombre el vocabulario ha de estar controlado.
- *Lenguajes controlados*. Son lenguajes controlados los lenguajes documentales propiamente dichos: tesauros, listas de encabezamientos de materia y clasificaciones. Presentan un vocabulario previamente elaborado, y admiten un limitado número de modificaciones en el momento de su utilización. El vocabulario controlado proporciona al usuario un punto de búsqueda, en vez de dos o más, y reduce la posibilidad de que la búsqueda sea incompleta. Sin embargo, puede perderse alguna información debido a la falta de especificidad y errores en el análisis.

Dependiendo de la coordinación de los términos, en función del momento en que se combinan los elementos que los componen, podemos distinguir:

- *Lenguajes precoordinados*. Combinan los términos en el momento de la descripción. Son lenguajes precoordinados las clasificaciones y las listas de encabezamientos de materia. En estos sistemas las distintas nociones o conceptos que se unen para expresar una materia o un tema ocupan un lugar determinado, es decir se introducen en el momento de la indización en un orden previamente establecido y la recuperación habrá de hacerse secuencialmente, siguiendo ese orden. Suelen ser muy precisos, pero también muy rígidos.
- *Lenguajes postcoordinados*. Combinan los términos en el momento de la recuperación. Son lenguajes postcoordinados los tesauros (lenguaje utilizado para la indización por descriptores), y si incluyéramos los lenguajes libres podemos mencionar también las listas de descriptores libres y las listas de

palabras clave. En estos sistemas las nociones o conceptos que se extraen en la indexación para expresar el tema o los temas del documento tienen todas la misma categoría y no se expresarán en ningún orden determinado. Serán los intereses de los usuarios en la fase de recuperación los que obliguen a la combinación de conceptos que se realizará utilizando los operadores booleanos, los operadores de comparación, etc.

Al combinar las características de control y coordinación es posible establecer cuatro grupos:

- *Lenguajes libres y precoordinados*: el sistema Precis.
- *Lenguajes controlados y precoordinados*: las clasificaciones y las listas de encabezamientos de materia.
- *Lenguajes libres y postcoordinados*: las listas de descriptores libres y las listas de palabras clave.
- *Lenguajes controlados y postcoordinados*: los tesauros.

Dependiendo de su estructura, los lenguajes documentales pueden ser:

- *Lenguajes de estructura jerárquica*. A la estructura jerárquica, también denominada arbórea o sistemática, responden determinadas clasificaciones (las clasificaciones jerárquicas): por ejemplo, la Dewey y la CDU. En ellas los conceptos se distribuyen en clases o categorías, donde dependen unos de otros. Estos lenguajes van de lo general a lo particular. Útiles en fondos enciclopédicos, resultan poco operativos para recuperar información, son sistemas muy rígidos y difíciles de actualizar, pueden servir para organizar colecciones.
- *Lenguajes de estructura asociativa o combinatoria*. Son aquellos que permiten la libre combinación de términos entre sí, eliminando la rigidez de los lenguajes de estructura jerárquica. A la estructura combinatoria o asociativa responden, por ejemplo, las listas de encabezamientos de materia y los tesauros, si bien es cierto que en la estructura de estos últimos participa también la estructura jerárquica. Estos lenguajes tienen una presentación alfabética y los conceptos se asocian o combinan libremente entre sí para expresar el tema o los temas.
- *Lenguajes de estructura sintáctica*. Surgen por la cantidad de ruido que producen los lenguajes combinatorios, que pueden poseer una red de relaciones semánticas pero no tienen gramática. La estructura sintáctica permite concretar y ordenar los conceptos. Este tipo de lenguaje requiere un cierto grado de automatización en los procesos tanto de indexación como de recuperación, por lo que su aparición está asociada al desarrollo tecnológico. Según su complejidad, se pueden diferenciar dos tipos: lenguajes sintácticos de gramática simple y lenguajes sintácticos de gramática elaborada. De

momento, los que aparecieron a finales de los años sesenta, como Syntol (Syntagmatic Organization Language) no han tenido demasiado éxito.

No corresponde definir todos y cada uno de los lenguajes documentales existentes en este tema, puesto que no habría ni tiempo ni espacio suficiente para hablar de todos ellos. De hecho, sería posible dedicar un tema a cada uno. Por ello, ofreceremos sólo una rápida visión de conjunto sobre los principales tipos de clasificación, los encabezamientos de materia, descriptores y tesauros.

Principales tipos de clasificación

Todos los días cada uno de nosotros realizamos continuamente el acto de clasificar, muchas veces sin darnos cuenta: clasificamos las personas como simpáticas y antipáticas, distinguimos las cosas que nos gustan de aquellas que nos disgustan, agrupamos objetos similares desde el punto de vista del uso, de la forma o del color, separamos los detergentes de los alimentos, etc., etc., etc. Desde Pitágoras la ciencia ya se clasificaba en Matemáticas, Aritmética, Geometría, Música y Astronomía, clasificación que se mantuvo durante más de dos milenios en el Quadrivium latino. Aristóteles afirmaba que “toda ciencia es o práctica, o productiva, o teórica”, al tiempo que dividía las ciencias teóricas en “física, matemática y metafísica”.

Se considera que las primeras clasificaciones bibliográficas adaptaron las antiguas clasificaciones de la ciencia. La clasificación más antigua que se conoce es la del templo de Horus en Egipto, y uno de los primeros intentos de clasificación bibliográfica se atribuye a Calímaco (260-240 a.C.), bibliotecario de la Biblioteca de Alejandría, aunque el primer sistema de clasificación bibliográfica como tal es el de Conrad Gessner en el siglo XVI. Sin embargo, habrá que esperar hasta mediados del siglo XIX para que grandes bibliotecarios norteamericanos como Cutter o Dewey establezcan las técnicas comunes para el tratamiento clasificatorio de los fondos de las bibliotecas.

Clasificar significa distinguir, reconocer las características que tienen en común los objetos o conceptos, agruparlos según estas características y, finalmente, separar los grupos así formados. Este concepto aplicado al saber que reside en los documentos ha permitido identificar contenedores ideales representados por las materias en las que el saber mismo puede ser dividido. El objetivo de la clasificación en el análisis documental consiste en identificar a qué clase o subclase se atribuye idealmente un documento y concretar este aspecto teórico en la elaboración de un catálogo sistemático de materia y/o en la ordenación física por materias de los documentos en las estanterías. Para poder realizar esto, la biblioteca tendrá que adoptar un sistema de clasificación que

proporcione una subdivisión precisa del saber, organizándolo en clases y subclases.

Los sistemas de clasificación son numerosos, y van desde el más rudimentario utilizado en los estantes de nuestra casa hasta aquellos que se basan en un fundamento científico y son adoptados institucionalmente tanto por las grandes bibliotecas nacionales como por pequeñas bibliotecas municipales. En cualquier caso, como lenguaje documental que representa de forma formalizada el contenido de los documentos, los sistemas de clasificación permiten la recuperación (manual o automatizada) de información requerida por los usuarios.

Los sistemas de clasificación, además de los requisitos de cualquier lenguaje documental, han de cumplir también las siguientes condiciones:

- no utilizan términos de lenguaje natural, sino signos o códigos normalizados basados en cifras, letras y otros símbolos gramaticales
- tienen que ser explícitos aunque concisos, es decir, que con el menor número de signos posibles expresen bien el contenido del documento
- siempre son lenguajes precoordinados
- completos, han de abarcar todos temas posibles en que se divide un área de conocimiento
- deben ser sistemáticos: proceder de lo general a lo particular, formando una estructura donde los conceptos se relacionan y ordenan en función de unas características específicas

Tipos de sistemas de clasificación

Se suelen distinguir distintos tipos de sistemas de clasificación dependiendo del criterio que se adopte, siendo los más frecuentes el contenido y la estructura.

Según su contenido o alcance pueden ser:

- *Clasificaciones enciclopédicas*: abarcan todas las ramas del saber. El más difundido y utilizado en España es la Clasificación Decimal Universal (CDU). Existen otros, entre los que se puede citar: Clasificación Dewey, Brunet, Cutter, Bliss o la LCC (Library of Congress Classification)
- *Clasificaciones especializadas*: profundizan en campos específicos del conocimiento, por lo que los términos utilizados corresponden a un área temática en particular. Algunas de las más conocidas clasificaciones de este tipo son la de la National Library of Medicine (EEUU), la Clasificación Decimal Astronáutica y la Clasificación de la OCDE.

Según su estructura pueden ser:

- *Clasificaciones monojerárquicas o jerárquicas*: también denominadas analítico-enumerativas. Sus clases, y subclases se organizan en forma arborescente, dando lugar una serie de divisiones y subdivisiones dependientes unas de otras. Los sistemas de estructura jerárquica pueden ser especializados, cuando sólo se ocupan de una rama del saber o enciclopédicos, cuando abarcan el universo de los conocimientos. Con frecuencia un sistema de clasificación especializado no es más que el desarrollo minucioso de una rama de un sistema enciclopédico aunque también pueden haberse creado expresamente para esa especialidad. No se pueden matizar las materias ni se pueden añadir nuevas fácilmente. Ejemplos serían la LCC o la Dewey.
- *Clasificaciones facetadas*: lenguajes documentales precoordinados, que a través de su organización permite construir áreas de conocimiento, creando un puente entre los sistemas jerárquicos y los asociativos, convirtiéndolo en un sistema analítico-sintético; lo cual significa que primero se analizan las distintas facetas relativas al documento y después, mediante síntesis, se llega a materializar su correspondiente número de clasificación. Clasificación menos rígida que las jerárquicas, adaptable fácilmente a la renovación de conceptos, se pueden añadir o modificar facetas sin perjudicar la estructura básica del sistema. Ejemplos serían la Colon Classification de Ranganathan o la de Vickery.
- *Clasificaciones híbridas*: también denominadas mixtas. Toman la estructura del conocimiento en grandes bloques de la estructura jerárquica pero incorporando las facetas, lo que da más precisión a la caracterización de los documentos. Enumera las materias básicas en tablas principales; las subdivisiones auxiliares comunes son aplicables a las materias básicas, neutralizando así la rigidez de la estructura jerárquica. El ejemplo clásico es la CDU.

Principales sistemas de clasificación

Entre los principales sistemas de clasificación utilizados en el mundo, podemos destacar:

- Clasificación expansiva de Cutter (1891). Este sistema fue formulado por Charles Ami Cutter. Dividió las materias en grandes grupos (7 tablas) que pueden ampliarse (de 8 a 27, según la necesidad de la biblioteca). Es una de las más antiguas e influyó sobre otras posteriores, especialmente sobre la Library of Congress Classification (LCC).
- Clasificación Decimal de Dewey (1873). Este sistema fue formulado por Melvil Dewey, para responder a las necesidades prácticas de un

determinado tipo de bibliotecas. La DDC es un sistema jerárquico, que se estructura en 10 grandes grupos temáticos que se subdividen de 10 en 10 y van de lo general a lo específico. La notación es numérica y consta de tablas principales, tablas auxiliares, subdivisiones generales y especiales y signos de correspondencia, además de los índices. Desarrollada y mantenida por la Decimal Classification Division de la Library of Congress, editada y comercializada por la OCLC, traducida a más de treinta lenguas distintas, va por su 22 edición y es sin lugar a dudas el sistema de clasificación más utilizado en el mundo en la actualidad. La Clasificación Decimal de Dewey es una clasificación bibliotecaria fundamentalmente enumerativa, que dió origen a la Clasificación Decimal Universal (CDU).

- Colon Classification (1933). Este sistema fue formulado por el matemático hindú Shiyali Ramamrita Ranganathan. Esta clasificación solamente se ha utilizado de forma general en India, sin embargo, ha ejercido gran influencia en los sistemas de clasificación e indización. Muchos autores ven este sistema de clasificación como un reflejo de la mentalidad oriental. La Colon Classification es el prototipo de las clasificaciones por facetas. Cada materia se puede descomponer en cinco grandes familias (PMEET): Personalidad, Materia, Energía, Espacio y Tiempo. La notación de este sistema es mixta y combina números arábigos, letras mayúsculas y minúsculas del alfabeto latino, algunas letras del alfabeto griego y signos de puntuación para conectar las facetas. La notación es laboriosa porque se confecciona para cada documento, pero está dotada de gran flexibilidad.

- Library of Congress Classification (1897-1948). La LCC es una adaptación del sistema expansivo de Cutter realizada en el siglo XIX por la necesidad de la Library of Congress de replantearse su sistema de organización de la colección debido al gran número de volúmenes. Es la colección más tradicional entre los esquemas enciclopédicos al uso. Al adaptar esta clasificación imperó el planteamiento pragmático, en ningún momento se planteó la organización sistemática del conocimiento, sino crear una herramienta capaz de almacenar y recuperar una colección con gran número de volúmenes. La notación se hace, por tanto, con una, dos o tres letras mayúsculas y un máximo de cuatro números arábigos. Las tablas que componen el sistema (21) tienen total independencia unas de otras, incluso pueden considerarse como clasificaciones individuales. Cada tabla incluye un índice detallado de materias pero no hay un índice general de todas las materias. Este sistema, a pesar de sus deficiencias (rigidez y ausencia de categorías generales), es muy apreciado por su excelente adaptabilidad en la organización de grandes colecciones de carácter enciclopédico (por ejemplo, bibliotecas universitarias) de libre acceso, por la regularidad de sus actualizaciones, pero sobre todo por la difusión de registros bibliográficos con dicha notación por parte de la LC. Es utilizada por más de un millar de

bibliotecas de todo EEUU, Australia y Reino Unido. En España, la primera implantación de la LCC fue en la Biblioteca de la Universitat Pompeu Fabra (1990) y después en la Biblioteca de la UJI (1992), aunque la satisfacción declarada por estas instituciones no ha servido para captar nuevos usuarios del sistema desde entonces.

La Clasificación Decimal Universal (CDU)

La Clasificación Decimal Universal es uno de los sistemas más extendidos en el mundo (124 países) tras la Clasificación Decimal de Dewey (135 países), y bastante más popular que la LCC. La CDU es la clasificación sistemática más difundida en España, aunque por supuesto no es la única utilizada. Según los resultados de una encuesta recientemente publicada (Aida Slavic, 2008) el número de instituciones que utilizan la CDU en el mundo se ha reducido últimamente, comparando los datos con los de los años 70 y 80, disminución que se acusa especialmente en Europa Occidental, Asia y América del Sur.

Este sistema parte del creado por Dewey en 1876, ya que los belgas Paul Otlet y Henry LaFontaine tradujeron y transformaron la Clasificación de Dewey con el propósito de organizar un catálogo de carácter internacional. La CDU nació así como sistema clasificatorio del Repertorio Bibliográfico Universal. Otlet y LaFontaine publicaron en 1905-1907 el Manuel du Répertoire bibliographique universel, basado en la 5ª edición de la DDC.

Sin embargo, la que ha quedado como edición oficial para el archivo básico de la CDU es la publicada por Otlet y LaFontaine, con la ayuda de Donker Duyvis, en 1927-1933 con el nombre de Classification Décimale Universelle.

Su puesta al día se realizó a través de la FID (Federación Internacional de Documentación), anualmente y por secciones, desde 1951 hasta su decadencia en los años 80s (la FID terminó desapareciendo en diciembre de 2000). Antes de desaparecer, en 1992, la FID transfirió la propiedad de la CDU así como su mantenimiento y distribución al consorcio formado por los editores que por entonces publicaban la CDU en holandés, francés, inglés, japonés y español, el UDC Consortium. Desde entonces, la gestión de este sistema de clasificación ha mejorado, con una revisión y actualización de más del 30% de las clases de la norma. Además, desde 1993 el sistema CDU se puede adquirir en forma de base de datos (en inglés), el UDC Master Reference File, que se actualiza el mes de enero de cada año. Actualmente existen varias ediciones de la CDU, impresas y electrónicas, en no menos de 39 idiomas.

Las características principales de la CDU son:

- Universal: abarca todo el saber humano.
- Internacional: sus notaciones numéricas la hacen inteligibles en cualquier país; el organismo que se ocupa de su actualización, la Asamblea CDU está

dentro de la Federación Internacional de Documentación, compuesta por representantes de todo el mundo.

- Jerárquica: de lo más general a lo más particular.
- Multidimensional: característica heredada de la Colon Classification: las divisiones analíticas permiten que un concepto aparezca en distintos lugares de la CDU según pertenezcan a un campo del saber determinado.

La CDU en España

Durante el siglo XIX, el sistema de clasificación más extendido en España fue el de Brunet. Procedente de Francia, donde gozó de una amplia difusión, este sistema dividía el saber en 5 grandes clases identificadas por letras y generaba 11.000 subdivisiones. Fue implementado en la Biblioteca Real (que posteriormente pasó a ser Biblioteca Nacional), en las bibliotecas públicas y populares y también en las bibliotecas universitarias. Este sistema fue el que predominó en España hasta la tercera década del siglo XX.

El primer difusor de la CDU en España fue Manuel Castillo, un bibliotecario de la Universidad de Salamanca. Tradujo la CDU al español y la publicó en 1895. El segundo intento de implementar la CDU en bibliotecas españolas fue en la Biblioteca de Ingenieros Militares, con el capitán Leopoldo Giménez, en 1908. Ricardo Codorníu publicó en 1911 una monografía sobre la Clasificación Decimal, con una traducción de los números de las tablas de selvicultura. El bibliotecario de la Biblioteca Provincial del Instituto General y Técnico de Badajoz, Ramón Gómez de Villafranca, realizó en 1911 una bibliografía utilizando la Clasificación Decimal. Antonio Paz y Meliá trató de implantar sin éxito la Clasificación Decimal en la Biblioteca Nacional. Al principio del siglo XX la implementación de la CDU en España eran casos aislados.

Comparado con el resto de España, la evolución de las bibliotecas de Cataluña fue diferente. En 1914, el bibliotecario Sebastián Farnés hizo una traducción diferente de la numeración de la CDU para asignar a la lengua catalana y Cataluña. Jordi Rubió i Balaguer fue profesor de clasificación y director de la Biblioteca de Cataluña. Fue el responsable de la primera traducción al catalán de la Clasificación Decimal en 1905.

En 1931 fue proclamada la II República en España, y durante ese período se desarrolló la política bibliotecaria, implementando la CDU en la Biblioteca Nacional, bibliotecas públicas, universitarias, municipales y otras, a pesar de que esa implementación no tenía un respaldo legal. El Gobierno de la II República creó algunas organizaciones para desarrollar las bibliotecas. La Junta elaboró normas e instrucciones sobre la organización de las bibliotecas por medio CDU. Textos como las "Instrucciones para el servicio de pequeñas bibliotecas" de María Moliner o las "Indicaciones para la organización de las Bibliotecas de los frentes, cuarteles y hospitales" de Teresa Andrés, publicados ambos en 1937, tuvieron una gran importancia por su originalidad y porque

fueron los primeros reglamentos de este tipo para dar forma a la aplicación de la CDU.

En 1939, tan sólo unos días después de finalizada la Guerra Civil, se estableció la obligatoriedad oficial del empleo de la CDU para la clasificación de los fondos de las bibliotecas españolas, según la Orden de 29 de julio de ese año. Se adoptó el "Sistema Decimal: Melvil Dewey modificado por el Instituto Internacional de Bibliografía de Bruselas y con las correcciones introducidas por su representación en Berlín", es decir, se adoptó la CDU (o el Sistema Decimal en su 3ª edición internacional, en lengua alemana) siguiendo a Alemania por la afinidad existente, entonces, entre España y Alemania. El bibliotecario Javier Lasso de la Vega fue quien instó al nuevo Gobierno para que legislara y posibilitara esta implantación oficial de la CDU. Redactó la Orden de 1939, debido a su colaboración con el nuevo Gobierno tras la Guerra Civil. Ello le llevó a tratar que se adoptara de forma oficial la CDU con las modificaciones emanadas de Berlín, omitiendo las ya publicadas ediciones internacionales y traducciones de la Clasificación Decimal que se habían realizado como la 4ª edición internacional en lengua inglesa de 1936 y la 5ª de Internacional de 1939 en lengua francesa. Sin embargo, la consumación de la implantación de la CDU se produjo con este decreto de 1939 que recogía una práctica bibliotecaria que había tenido plena vigencia durante la última década.

Finalizada la Guerra Civil desapareció el Gobierno autónomo de la Generalidad y la Biblioteca de Cataluña de nuevo pasó a depender de la Diputación de Barcelona, al igual que había ocurrido durante el período de Primo de Rivera. La nueva Administración Central marcó las directrices de las bibliotecas catalanas, y mediante el citado decreto de 29 de julio de 1939 quedaba implantada la versión de 1934 de la Clasificación Decimal Dewey. Pero en Cataluña desde hacía más de una década el Sistema Decimal ya se había implantado, de forma total, en la red de Bibliotecas Populares a instancia de Jordi Rubio i Balaguer. En la Biblioteca de Cataluña se inició la implantación más tarde, mientras se producía el traslado de ésta desde su anterior ubicación en el Palacio de la Diputación, hasta el nuevo edificio del Hospital de Santa Cruz y San Pablo en el que continúa en la actualidad. Así las nuevas directrices en la recién instalada Biblioteca hicieron que en ella se implantara finalmente la CDU y no la Clasificación Decimal, con las modificaciones realizadas en 1905 por el Instituto Internacional de Bibliografía. Las Bibliotecas Populares catalanas, al igual que las restantes del Estado español, quedaron sujetas a esta organización de los libros en salas y depósitos, aunque en algunos casos en los depósitos no quedó totalmente implantada, pues en muchas bibliotecas no existió el acceso directo a los estantes por parte de los usuarios. Sin embargo, sí se hizo uso de la CDU para la ordenación y distribución de los catálogos sistemáticos.

En la actualidad, mediante el decreto de 19 de mayo de 1989, por el que se aprueba el Reglamento de Bibliotecas Públicas del Estado y el Sistema Español de Bibliotecas (vigente transitoriamente según se establece en la Ley 10/2007, de 22 de junio, de la lectura, del libro y de las bibliotecas) se establecen normas reglamentarias de organización y funcionamiento de las Bibliotecas de titularidad estatal. En lo que hace referencia al tratamiento técnico de los fondos queda expresada la derogación de la "Orden de 29 de julio por la que se implanta el sistema bibliográfico decimal en la clasificación de los fondos de las Bibliotecas Públicas del Estado" (disposición derogatoria 1). Por tanto, a través de esta disposición queda derogado el empleo obligatorio, por parte de los bibliotecarios españoles, de la CDU no sólo en lo que hace referencia a la ordenación de los fondos, sino también abarca a la organización de los catálogos, no disponiendo obligatoriedad de emplear sistema clasificatorio alguno.

Además, con la creación de los Estatutos de las Comunidades Autónomas, entre 1979 y 1983 se han creado competencias específicas en materia bibliotecaria en cada una de ellas. No obstante las normativas existentes, por el momento no recogen reglamentación relativa a la clasificación a excepción del Reglamento interno de organización de los servicios de las bibliotecas públicas de titularidad estatal gestionadas por la Comunidad de Castilla y León (Decreto 263/1997, de 26 diciembre 1997). De todo ello se infiere que en la actualidad existe una "laguna legal" respecto del sistema clasificatorio a emplear, y el nuevo órgano creado mediante el art. 15 de la Ley de la lectura, del libro y de las bibliotecas (2007), el Consejo de Cooperación de Bibliotecaria, órgano colegiado de composición interadministrativa que canalizará la cooperación bibliotecaria de forma voluntaria entre las administraciones públicas, regulado mediante RD 1573/2007, de 30 de noviembre, sustituto del poco fructífero Consejo Coordinador de Bibliotecas creado en 1989, ya no está facultado para dictar normas sino para simplemente promover planes específicos que mejoren las condiciones de las bibliotecas y sus servicios entre los que se incluirá la "adopción de estándares e indicadores que faciliten un servicio público de calidad", planes que "se evaluarán y actualizarán periódicamente y adoptarán como marco de referencia las directrices, pautas, estándares, recomendaciones, normas técnicas u otros documentos similares emanados de organismos nacionales e internacionales que les sean de aplicación".

En conclusión, en estos momentos de transición hacia nuevas configuraciones teóricas y prácticas en el ámbito de la clasificación documental y de globalización en el tratamiento de la información, podemos decir que en el sistema bibliotecario español las puertas están abiertas para un cambio en el sistema de clasificación.

Ediciones de la CDU en español

AENOR (Asociación Española de Normalización), una entidad privada sin ánimo de lucro, designada legislativamente como organismo normalizador y certificador en España, reconocida internacionalmente en el desarrollo de la normalización, es miembro fundador del UDC Consortium y responsable de la edición en español. De esta forma, en España, la CDU adquiere rango de norma UNE 50001:2000 (condición que no posee, por ejemplo, en la British Standard Institution (BSI): BIP 0017:2006 no es una BS, es una simple publicación).

La CDU se publica en soporte impreso y en CD-ROM. La CDU impresa se presenta en tres volúmenes: el primero contiene las tablas sistemáticas 0/5 así como las tablas auxiliares completas cuyos números permiten expresar facetas adicionales de la materia; el segundo las tablas sistemáticas 6/9; y el tercero un índice temático, por orden alfabético, de los conceptos incluidos en la CDU, con indicación del término y la notación correspondiente al concepto de que se trate. Existe también una edición abreviada de la CDU, impresa, adaptación de la última edición de la Norma UNE 50001:2000, cuyo objetivo de facilitar las tareas de clasificación de obras bibliográficas. Así, únicamente se han eliminado aquellos números clasificatorios menos usuales y característicos en el contexto científico y cultural de los países de habla hispana, sin que se haya alterado la estructura, sistematización y notación general de las tablas.

Por último, también existe una edición electrónica en español de la CDU, que recoge en CD-ROM el contenido íntegro de la CDU impresa, y cuenta con un potente motor de búsqueda para la consulta de su contenido. El software utilizado para el desarrollo de la CDU en soporte informático es el de hipertexto, sistema de seguimiento, tanto de texto como de gráficos, que permite ir de un documento o apartado a otro sin tener que visualizar todas sus páginas.

Estructura de la CDU

Este sistema agrupa el conocimiento en 9 clases, dejando la clase 4 vacía, y siguiendo el principio de subordinación lógica, a semejanza del sistema Dewey, cada uno de estos grupos puede subdividirse a su vez, y así sucesivamente descendiendo infinitamente en el nivel jerárquico. Es una clasificación decimal, puesto que los números se contemplan como si fueran fracciones decimales, de manera que cuanto más largo es el número, más específica es la materia.

En cuanto a la temática que abarca, tiene vocación universal y enciclopédica, la CDU se compone de tablas principales, tablas auxiliares (comunes y especiales), y un índice alfabético que remite a las tablas principales.

Encabezamientos de materia

Un encabezamiento de materia es la expresión lingüística (palabra o conjunto de palabras) que representa el contenido temático de un documento y que se utiliza para hacer búsquedas en un catálogo, bibliografía o índice. Puede constar de un elemento, si el tema se expresa con un solo término, o de varios elementos, si el tema se expresa mediante la combinación de un conjunto de términos dispuestos en una cadena según unas reglas sintácticas preestablecidas. Esta sintaxis especial es la que precisamente diferencia los encabezamientos de materia de los descriptores.

En el contexto de los lenguajes documentales, las listas de encabezamientos de materia (LEM) son lenguajes controlados precoordinados, formados por un conjunto de términos extraídos del lenguaje natural.

El objetivo principal de una lista de encabezamientos de materia, al igual que cualquier lenguaje documental controlado, es promover la utilización de unas prácticas indización normalizadas dentro de una institución o entre diferentes instituciones o entre diferentes servicios de indización, principalmente si estas instituciones o servicios intercambian o comparten la catalogación o la información bibliográfica. Estas listas se usan preferentemente en las bibliotecas nacionales, universitarias, públicas y municipales, y en aquellas bibliotecas o centros de documentación cuyos fondos son esencialmente enciclopédicos (es decir, no se limitan a un dominio especializado del conocimiento). Aunque desde el punto de vista teórico los centros especializados podrían utilizar listas de encabezamientos, en la práctica suelen utilizar tesauros.

La catalogación alfabética de materias, además, no se aplica a todo tipo de documentos: no son objeto de entradas las obras de creación, las literarias y las clásicas, porque se considera que su contenido intelectual resulta tan heterogéneo y complejo que sería imposible reducirlo a un simple encabezamiento de materia. Se supone que se localizan a través del catálogo alfabético de autores o a través del título. Sólo se hace una excepción: las antologías o repertorios de obras literarias sí que son objeto de indización mediante encabezamiento de materia.

Principios básicos de la indización mediante LEM

- Especificidad. Cuando se utiliza una lista de encabezamientos de materia es necesario seleccionar el término existente más específico para representar una determinada noción. La especificidad está relacionada con la exactitud mediante la cual una noción concreta del documento se encuentra representada por un término de indización. Cuando una noción concreta se encuentra representada por un término más general se produce una pérdida de especificidad. Se han de identificar las nociones de la forma

más específica posible. Pero, en algunos casos, puede resultar más práctico emplear una noción más general como, por ejemplo:

- a) cuando el indizador considera que un exceso de especificidad podría tener un efecto negativo sobre el funcionamiento del sistema de indización; o también
- b) en función del peso asignado a una determinada noción por el autor.

Si el indizador considera que una idea no se encuentra lo suficiente desarrollada, o que sólo se trata de una alusión fortuita del autor, entonces puede quedar justificada la selección de un término de indización más general.

- **Univocidad.** Un catálogo alfabético de materias debe reunir bajo un sólo encabezamiento todos los documentos que traten de manera exclusiva o principal ese mismo tema o materia. Se tratan como sinónimos los términos cuasisinónimos, las variantes ortográficas, las variedades lingüísticas, las abreviaturas y los acrónimos, las formas alternativas de los nombres propios o términos equivalentes en otra lengua (cuando se considera que el uso del término extranjero se encuentra razonablemente extendido), así como la forma permutada de algunos encabezamientos compuestos (es decir, encabezamientos seguidos de subencabezamientos o subdivisiones propias de aquel encabezamiento concreto), y también los términos científicos o técnicos (cuando se admite como encabezamiento aceptado el término popular) y, por último, los términos obsoletos.
- **Adecuación al uso lingüístico.** La lista de encabezamientos de materia debe representar el uso común de los términos en el lenguaje natural, para evitar las dificultades que genera el empleo en jergas muy especializadas, préstamos lingüísticos o extranjerismos (salvo que no exista equivalente). También debe seguir las reglas gramaticales que se determinan en cada lengua y, por supuesto, adecuarse a la lengua de los usuarios, que se supone que es la lengua del catálogo. La indización multilingüe es posible, pero requiere el mantenimiento por separado de las listas de encabezamientos según la lengua, siendo posible establecer relaciones entre cada una de las listas.

Principios generales para la asignación de los encabezamientos

Existe una regla general según la cual se asigna a la obra que se está catalogando el encabezamiento o encabezamientos que expresan mejor el contenido temático o la forma del documento con la intención de facilitar el acceso al tema o temas principales. Para poder asignar un encabezamiento a un tema determinado, el tema tratado ha de ocupar, como mínimo, el 20% de la obra.

- Número de encabezamientos: El número de encabezamientos necesarios varía en función de la obra que se está catalogando. A veces, no hace falta más que un encabezamiento solo. Por regla general, un máximo de seis resulta apropiado. Si se asigna más de un encabezamiento, es mejor citarlos por orden de relevancia. Un tema ha de estar representado por el encabezamiento que lo exprese con más exactitud y precisión. Es decir, el encabezamiento no debe ser más general que el tema tratado. En muchos casos, la especificidad se consigue mediante la asignación de un encabezamiento principal formado por una palabra o un grupo de palabras. Otras veces, la especificidad se consigue mediante el uso de subdivisiones.
- Profundidad de la indización: Un determinado encabezamiento, dependiendo del lugar que ocupa dentro su jerarquía, puede comprender varios subtemas que también tienen sus respectivos encabezamientos en la lista de autoridades. En este caso, sólo se asigna el encabezamiento que designa más exactamente el contenido general de la obra. No se han de asignar los encabezamientos que, normalmente, representan los subtemas y que se encuentran comprendidos en la cobertura conceptual del encabezamiento en cuestión. Por ejemplo, el título “Iniciación a la gimnasia” se indexa por “Gimnàstica” [Es decir, no se han de asignar encabezamientos separados para barras de equilibrio, barras paralelas, barras transversales, potro, anillas, etc., en lugar de, o además del encabezamiento “Gimnàstica”]
- Reglas de asignación del número de encabezamientos: Para saber el número de encabezamientos que hemos de asignar a un determinado documento existen una serie de principios orientativos que nos pueden ayudar a clarificar esta cuestión. Este principios son:
 - Dos o tres encabezamientos relacionados. Lo cual significa que si existe un encabezamiento que representa los dos o tres temas tratados en la obra y, dentro de su ámbito conceptual, no incluye también otros temas, se asigna este encabezamiento en lugar de los dos o tres encabezamientos más concretos. Por ejemplo, el título “Historia del transporte terrestre, marítimo y aéreo”, se indiza por: “Transport—Història”
 - Regla de los tres. Se refiere al hecho que si un tema general abarca más de tres subtemas, pero la obra que se está catalogando sólo trata dos o tres de estos subtemas, es mejor asignar los dos o tres encabezamientos apropiados que el encabezamiento más general. Por ejemplo, el título “Fructicultura en Cataluña, 1999: melocotones y manzanas” se indiza por “Préssecs--Catalunya—Estadístiques” y “Pomes-Catalunya—Estadístiques”. Si la obra trata de más de tres subtemas, se asigna el encabezamiento general, a menos que la regla de los cuatro sea aplicable.

- Regla de los cuatro. Afirma que, en algunos casos, es preferible asignar encabezamientos diferentes para representar los cuatro subtemas de un concepto muy amplio. Si un encabezamiento es muy amplio y cada uno de los subtemas tratados sólo constituye una pequeña parte de este concepto, se asignan los cuatro encabezamientos. Nunca se deben superar los cuatro encabezamientos. Por ejemplo, si una obra trata de cuatro escritores catalanes, se puede asignar un encabezamiento para cada uno de ellos, puesto que el encabezamiento “Literatura catalana--Història i crítica” abarcaría todos el escritores catalanes.

- Temas con elementos múltiples. Se aplica cuando una obra trata de un tema complejo para el que no existe ni se puede crear un único encabezamiento. En este caso concreto, se asignan múltiples encabezamientos para destacar los diferentes aspectos del tema. Por ejemplo, el título “Morbosidad cancerígena y mortalidad entre los cerveceros daneses” se indiza por “Càncer—Dinamarca”, “Càncer—Dinamarca--Mortalitat”, “Cervesers--Malalties--Dinamarca” y “Cervesers—Dinamarca--Mortalitat”.

- Principio de la objetividad: Se ha de evitar utilizar encabezamientos que expresen juicios de valor u opiniones personales sobre temas o materiales. Sin embargo, sí que se puede tener en cuenta la intención del autor o editor sobre el tipo de lector a quien se dirige la obra, el tratamiento como hecho verídico o de ficción, etc.

Tipos de encabezamientos

Es posible distinguir los encabezamientos de materia según su contenido, según su estructura formal y según su posición dentro de la estructura del encabezamiento

Según su contenido, distinguimos:

- Encabezamientos temáticos y de forma. Un encabezamiento temático es aquel que representa un concepto o un objeto tratado en un documento e indica el contenido temático de este documento. El encabezamiento de forma indica principalmente lo que el documento es, aunque, a veces, también designa el tema tratado por el documento. Por ejemplo, “Enciclopèdies i diccionaris catalans” es un encabezamiento de forma que puede indicar que el documento es una enciclopedia o bien que el documento es un tratado sobre las enciclopedias escritas en catalán.
- Encabezamientos de nombre personal, de entidad y títulos. Normalmente no suelen aparecer en las listas, salvo aquellos que se usan como modelo, como ejemplos, o que llevan subdivisiones temáticas especiales o notas explicativas. También se incluyen una serie de encabezamientos ambiguos

como son los nombres de los personajes bíblicos, clanes, casas nobiliàries, dinastías, familias y casas reales, nombres de personajes de ficción, legendarios y mitológicos o dioses y diosas. Las formas correctas de los nombres personales, de los nombres de entidad, de jurisdicciones y los títulos uniformes se establecen de acuerdo con las reglas de catalogación utilizadas y adoptan la misma forma que presentarían en el catálogo de nombres y de títulos.

- Encabezamientos geográficos. Existen dos tipos de encabezamientos geográficos: nombres geográficos de jurisdicciones políticas y administrativas y nombres geográficos no jurisdiccionales. Los nombres geográficos de jurisdicciones políticas y administrativas (como nombres de países, estados, autonomías, departamentos, provincias, diócesis, ciudades, etc.) sólo constan en la lista cuando se necesita añadir una subdivisión temática o cronológica, como por ejemplo.: “Catalunya--Història--1213-1276, Jaume I”. Los nombres geográficos jurisdiccionales se construyen según el capítulo correspondiente de las reglas de catalogación. Los nombres geográficos jurisdiccionales reciben un tratamiento diferente si se utilizan en la catalogación descriptiva o en la catalogación de materias. En la catalogación descriptiva se usa, como punto de acceso, el nombre geográfico jurisdiccional propio de la época que contempla la obra, mientras que en la catalogación de materias, cuando el nombre del país, estado, ciudad, etc. ha cambiado, sin afectar substancialmente el área geográfica cubierta por la jurisdicción, como punto de acceso de materias sólo se usa el último nombre, sin tener cuenta la época contemplada por el documento. Por ejemplo.: “Sri Lanka” [y no] “Ceilan”. Los nombres geográficos de jurisdicciones políticas y administrativas que han cambiado de nombre y que no abarcan la misma área geográfica se pueden usar también como punto de acceso de materia. Por ejemplo: “Lleó (Regne)”.

Desde el punto de vista de la estructura formal, los encabezamientos de materia pueden ser:

- Encabezamientos simples o de un solo término formados por un sustantivo o un adjetivo sustantivado. Los nombres que representan objetos concretos generalmente se citan en plural y los nombres que representan conceptos abstractos aparecen en singular, ej.: “Gossos”; “Filosofia”.
- Encabezamientos formados por un sustantivo seguido de un adjetivo, ej.: “Literatura barroca”
- Encabezamientos formados por dos sustantivos unidos por una conjunción que normalmente representan temas o conceptos relacionados, encontrados o que establecen una relación de causa y efecto, ej. “Església i Estat”

- Encabezamientos formados por más de un sustantivo unido por preposición que se pueden referir a un único concepto o que establecen una relación entre conceptos diferentes, ej.: “Libres de viatges”
- Encabezamientos invertidos que sitúan en segundo término una palabra poco significativa, ej.: “Galois, Teoria de”
- Encabezamientos seguidos de paréntesis que sirven por distinguir dos términos homónimos, por precisar el contexto en el cual se utiliza un término, por aclarar términos ambiguos o por indicar una localización geográfica. Las aclaraciones entre paréntesis generalmente corresponden a disciplinas, categorías o tipos de objetos. Ej., “Canvi (Psicologia)”, “Canvi (Filosofia)”, “Canvi (Permuta)”.
- Encabezamientos compuestos formados por un encabezamiento y un subencabezamiento o subdivisión. Las subdivisiones funcionan como una extensión del encabezamiento principal. Normalmente, expresan aspectos de este encabezamiento. El concepto reflejado por la subdivisión puede ser temático, de forma, geográfico o cronológico. Ej.: “Sals--Efectes fisiològics”

Según su posición dentro de la estructura del encabezamiento distinguiremos entre:

- Encabezamiento: elemento principal que representa el contenido temático de un documento
- Subencabezamiento: cuando se utilizan, elemento acompaña de forma subordinada a un encabezamiento.

Subdivisiones bajo nombres geográficos, nombres de lengua y nombres comunes, o de aplicación libre

- Subdivisiones comunes. Por subdivisión común o de aplicación libre (Free-Floating subdivisions) se entiende aquella subdivisión temática o de forma que el indizador asigna debajo determinados temas sin que su aplicación haya sido editorialmente establecida y sin que aparezca recogida en la lista de autoridades bajo cada uno de los encabezamientos de la misma categoría.
- Subdivisiones geográficas. Los encabezamientos que llevan la indicación que se pueden subdividir geográficamente, se les puede aplicar la subdivisión geográfica directa (ej.: “Biblioteques--París” (LEMAC)) o indirecta, interponiendo el nombre del país o de la unidad jurisdiccional de primer orden delante del más pequeño (ej.: Libraries--France--Paris (LCSH))
- Subdivisiones cronológicas. Es necesario destacar que las subdivisiones cronológicas, en contra de lo que muchos indizadores piensan, no son de

aplicación libre, sino que se establecen editorialmente a partir de la lista. La política en los catálogos de materia de permitir simultanear las subdivisiones de aplicación libre con las subdivisiones cronológicas preestablecidas ha provocado una gran dispersión de los encabezamientos de materia que aparecen en los catálogos, puesto que años contiguos o fechas próximas expresan, muchas veces, una misma época o un mismo periodo. La aplicación de las subdivisiones de tiempos preestablecidas implica que una misma obra llevará tantos encabezamientos como sea necesario con tal de expresar las diferentes etapas cronológicas. De esta forma, una antología de poesía catalana de los siglos XIX y XX llevará dos encabezamientos de materia diferentes: "Poesía catalana--S.XIX" y "Poesía catalana--S.XX". Los periodos de tiempos también se expresan mediante un único encabezamiento, un encabezamiento seguido de un adjetivo o mediante un encabezamiento compuesto en el cual la subdivisión que se aplica directamente bajo el encabezamiento principal es una subdivisión cronológica, ej.: "Renaixement"; "Pintura barroca"; "Segle XVIII"; "Xilografia--S. XVIII"; "Música-- 500-1400".

- Subdivisiones de género o forma. Estas subdivisiones no designan la materia propiamente dicha, pero tampoco su soporte. Es decir, no son una mención en la materia de la designación general o específica del material, a pesar de que se puede encontrar este uso impropio en algunos catálogos (Microfitxa, CD-ROM, DVD). Estas subdivisiones se utilizan para reflejar la configuración del discurso que trata una materia, yendo desde aspectos externos (ej. "Fotografies") hasta el mismo género (ej. "Bibliografia"), pasando por muchos estados intermedios (ej. "Revistes", "Diccionaris", "Congressos"). Hay cierta confusión respecto a su uso, porque además algunas de estas subdivisiones pueden ser también encabezamientos de materia (ej. "Bibliografia").

Orden de citación de las subdivisiones y significado de los elementos

Un encabezamiento principal puede ir seguido de una o más subdivisiones del mismo o de diferente tipo.

El orden más usual de citación de las subdivisiones responde al modelo [Lugar] \$x [Tema] \$y [Periodo cronológico] \$v [Forma], por ejemplo.: Catalunya |x Condicions socials |y S.XX |v Bibliografia.

Otro orden de citación bastante frecuente es: [Tema] \$z [Lugar] \$x [Tema] \$y [Periodo cronológico] \$v [Forma], por ejemplo.: Ferrocarrils |z França |x Vies |x Història |y S.XIX |v Llibres de làmines; y [Tema] \$x [Tema] \$z [Lugar] \$y [Periodo cronológico] \$v [Forma], por ejemplo.: Art |x Col·leccions privades |z Itàlia |x Història |y S.XVI |v Bibliografia

La subdivisión cronológica y de forma (si la tiene) van siempre al final, en este orden, siendo repetible sólo la de forma. La subdivisión geográfica tiene, sin embargo, cierta complejidad en lo que se refiere a su posición: cuando es el encabezamiento, la subdivisión temática se pospone. En encabezamientos compuestos en los que bien el encabezamiento temático, la subdivisión temática o la subdivisión común (si la tiene) admiten la subdivisión geográfica, alguno o todos, diremos que ésta va en la posición permitida más distante del encabezamiento temático. Por ejemplo, el encabezamiento "Aigua" se puede subdividir geográficamente, "Aigua--Catalunya", pero si se le aplica la subdivisión común "Aspectes econòmics", que también se puede subdividir geográficamente, el orden de citación correcto será: "Aigua--Aspectes econòmics--Catalunya".

El significado de los elementos de un determinado encabezamiento se puede verificar construyendo una frase en la cual los elementos individuales de la cadena sean enunciados en orden inverso, es decir, empezando por el final (exceptuando la subdivisión cronológica, que se suele formular siempre al final). Así un encabezamiento como "Literatura alemanya--Història i crítica--Revistes" se puede leer como [Una revista sobre la historia y crítica de la literatura alemana] o "França--Relacions exteriors--1792-1815--Fonts--Bibliografia" se puede leer como [Una bibliografía de fuentes para las relaciones exteriores de Francia durante el periodo que abarca del 1792 al 1815].

Tesauros y descriptores

La explosión de la información, su diversificación y especialización, desbordaron la capacidad de los lenguajes clasificatorios para representar y recuperar el contenido de los documentos. La aparición del sistema Unitérmino de Taube (1951-1953), la elaboración del índice KWIC de Luhn (1959) y la indización por conceptos (descriptores) de Mores (1960) revolucionaron las técnicas de indización y pusieron las bases teóricas y metodológicas para la creación en los años 60 de los primeros tesauros.

Definición de tesoro

A partir de la norma UNE 50106:1990, Directrices para la creación y desarrollo de tesauros (equivalente de la ISO 2788:1986) los tesauros se pueden definir según su función y según su estructura:

- Por su función, se pueden definir como instrumentos de control terminológico. Controlar el vocabulario significa identificar dentro de un campo semántico todos los conceptos que son representados por más de un término. La identificación de términos equivalentes hace posible minimizar la

pérdida de información en las búsquedas realizadas en un sistema documental automatizado.

- Por su estructura, los tesauros permiten conocer todos los términos relacionados con un concepto determinado, lo que ayuda a añadir más términos adecuados para enriquecer tanto los análisis de contenido de los documentos como las estrategias de búsqueda para recuperar información.

El tesoro es un prototipo de lenguaje de indización y recuperación controlado, que se basa en la postcoordinación de sus descriptores. Además representa de manera unívoca los conceptos de los documentos evitando así los problemas de homonimia, sinonimia, polisemia (relaciones de significado), así como el establecimiento de las relaciones jerárquicas y de relación entre los descriptores que lo componen.

Los descriptores son términos del lenguaje natural que por medio de un proceso de selección llega a formar parte del vocabulario de un tesoro. La condición que debe superar un término para adquirir la condición de descriptor es que sea el más representativo del concepto que se quiere representar dentro del tesoro. Con la elección de este término como único representante válido de entre otros muchos con significados idénticos o casi idénticos se consigue la univocidad del lenguaje documental, en este caso del tesoro. Un descriptor, por lo tanto, es un término que representa de manera unívoca un concepto dentro de un lenguaje documental.

Las palabras clave o los unitérminos se diferencian de los descriptores en que, a pesar de compartir la procedencia del lenguaje natural y de tener la función de representar conceptos, las palabras clave carecen del control estricto de la sinonimia, de las relaciones asociativas y, mucho menos de las jerarquías. Los descriptores se caracterizan por presentar unas relaciones semánticas y mayor control que las palabras clave.

Elementos de un tesoro

Un tesoro se compone esencialmente de descriptores y no descriptores

- Descriptores: Son palabras o expresiones incluidas en un tesoro y escogidas del lenguaje natural entre un conjunto de términos equivalentes para representar sin ambigüedad una noción contenida en un documento o para formular una búsqueda documental.
- No descriptores: Llamados también términos equivalentes o no preferentes. Son sinónimos o cuasi-sinónimos de los descriptores o términos que designan en el lenguaje natural conceptos afines a los que cubren los descriptores. Los no descriptores no pueden ser utilizados para la indización del documento o para formular búsquedas, pero cada uno de ellos reenvía a uno o dos descriptores para representar los conceptos correspondientes, por lo que aumentan la tasa de precisión. Estos términos son puntos de acceso

no aceptados que facilitan el paso del lenguaje natural al lenguaje controlado, permitiendo la elección de descriptores pertinentes.

Relaciones entre los descriptores

Entre los descriptores se establecen cuatro tipos de relaciones fundamentales:

- *Asociativas*: es el tipo de relación más importante del tesoro por ser la que le define como lenguaje de estructura combinatoria, hecho que le diferencia de otros. Es una relación recíproca que identifica descriptores afines, que no jerárquicos entre sí, sean o no de una misma familia, en una dirección horizontal. Todo descriptor relacionado asociativamente con otro implica que éste, a su vez, mantiene idéntica relación con aquel. Utiliza la notación TR (Término relacionado), en inglés RT (Related Term) Ej.:

Etnolingüística

Etnografía

TR Etnografía

TR Etnolingüística

- *Jerárquicas*: define los descriptores más genéricos y más específicos, con sus correspondientes clases, especies, familias, partes, etc. A diferencia de los sistemas jerárquicos tradicionales, esta relación se establece en función del resto de las relaciones de un concepto, sin las cuales la jerarquía no siempre sería posible: un descriptor puede tener términos más genéricos o más específicos relacionados con él porque a su vez cuenta con relaciones asociativas, de sinonimia o definitorias que le permiten tener cualquiera de las tres relaciones restantes, sin que al mismo tiempo, ninguna sea excluyente de otras. Las jerarquías de un tesoro son además, mucho más completas porque pueden tener varias escalas. Es una relación recíproca, de modo que un mismo término específico de uno genérico guarda con este una relación inversa. La notación empleada es TG (Término Genérico), en inglés BT (Broad Term) y TE (Término Específico), en inglés NT (Narrow Term). Ej:

Etnolingüística

Lingüística

TG Lingüística

TE Etnolingüística

Sociolingüística

- *De equivalencia*: también llamadas relaciones de sustitución o preferenciales. Dirigen desde el sinónimo, cuasisinónimo o antónimo donde se centra la demanda, hacia el descriptor autorizado o término preferencial. Esta relación supone una traducción del término natural (libre) al artificial (normalizado). Las notaciones de esa relación son: USE y UP (Usado por) o en inglés USE y UF (Use For). USE remite de un término no admitido en el tesoro a un descriptor admitido y UP informa de los términos sinónimos, polisémicos, etc., no admitidos de un descriptor. Ej.:

Contador

Contable

USE Contable

UP Contador

- *Definitorias o aclaratorias*: se utilizan para delimitar la aplicación de un descriptor. Puede describir el significado por el cual se va a usar un término con sinonimia, acotar un significado a una única opción excluyendo algún sentido, eliminar ambigüedades, desarrollar siglas o abreviaturas, etc. Se utiliza con las notaciones NA (Nota de Alcance), en inglés SN (Scope Note). Pueden ser de varios tipos: definición completa; definición restringida; inclusión de términos en relación jerárquica con el descriptor; descripción de cómo debe usarse el término en la indización; e historia del descriptor.

Formas de presentación de los tesauros

Los términos que componen un tesauro y sus relaciones pueden presentarse de varias formas:

- *Alfabética*: contiene descriptores y no descriptores ordenados alfabéticamente. Con cada descriptor aparece el conjunto de sus relaciones semánticas (indicación del campo semántico al que pertenece, notas aclaratorias, equivalencias semánticas, descriptores genéricos, específicos y asociados). Ej. Agrovoc (tesauro agrícola multilingüe) (3ª ed., Organización de la Naciones Unidas para la Agricultura y la Alimentación-FAO, Roma, 1995). Cab (Commonwealth Agricultural Bureaux, Wallingford-Oxon, UK) Materias: Agricultura y áreas relacionadas (silvicultura, nutrición y parasitología, desarrollo rural, ingeniería agrícola, producción de semillas, etc.)
- *Sistemática*: contiene los descriptores estructurados en función de categorías. Los descriptores se sitúan en sus áreas contextuales donde son jerarquizados por categoría semántica y dotados de distintos tipos de relaciones. El establecimiento de categorías se puede realizar a partir de facetas o de campos temáticos. Ej.: Unesco 1ª ed. Multidisciplinar y versiones en varios idiomas, a partir del original inglés. Fuentes terminológicas: está basado en parte en el Sistema General de Clasificación-BS Asis (2ª ed., American Society for Information Science-ASIS, 1998) Tesauro de Biblioteconomía y Documentación y áreas afines. Es un tesauro sistemático con facetas pero con una concepción original. Tesauro de educación de unesco-oie (5ª ed., 1991, versión española, UNESCO) Tesauro utilizado en las bases de datos de la Oficina Internacional de Educación (OIE) y otros centros educativos.
- *Gráfica*: muestras sobre todo las relaciones jerárquicas existentes entre los descriptores y los reenvíos hacia términos asociados. Pueden darse en estructura arborescente, diagrama de flechas y terminogramas. Ej.: Eurovoc

1ª ed. (terminogramas) Tesouro europeo de la educación (terminogramas) (3ª ed., Luxemburgo, Oficina de Publicaciones Oficiales de las Comunidades Europeas, 1991) Es un tesouro multilingüe de educación, con presentación gráfica, en nueve idiomas comunitarios (alemán, danés, español, francés, griego, inglés, italiano, neerlandés y portugués)