

Normalización terminológica para la sistematización de la información patrimonial

Antonio Martín Pradas y María José Fitz Canca, Centro de Documentación y Estudios del IAPH

INTRODUCCIÓN

Desde su origen, el ser humano ha ido poniendo nombre a los elementos que formaban parte de su contexto físico y simbólico, designando cada cosa por su nombre. Ante el volumen de palabras que designaban todo tipo de objetos, animados e inanimados, se han ido creando diccionarios, en los que se especifica el significado y la acepción de cada una de las palabras que designan los objetos, elaborándose grandes enciclopedias, muy usadas a lo largo de siglos.

Para organizar los términos han surgido clasificaciones que estructuran el conocimiento de las distintas ramas del saber. De esta forma se han gestado los tesauros, como compendios de conocimientos y terminología de una disciplina específica.

Con la aparición de las nuevas tecnologías y la creación de las bases de datos para controlar de forma asequible la información de un conjunto de objetos, se observó la necesidad de normalizar la terminología utilizada con el objetivo de que todas las personas que usaran las aplicaciones hablasen el mismo idioma, o mejor dicho, que usasen las mismas pala-

bras para catalogar objetos semejantes, lo que permite una eficaz recuperación de la información.

Hoy día la normalización terminológica es el pilar central de los sistemas de gestión documental, a través de la cual quienes catalogan han de extraer la información de cada objeto para ponerla al servicio de un determinado colectivo especializado. Del mismo modo, deben controlar el análisis documental, que consiste en estudiar los objetos a través de una serie de pasos que permiten representarlo terminológicamente a través de la indización, lo que ayuda a la formación de unos listados que facilitan su localización.

El producto resultante de la indización son las listas de términos alfabéticos, en lenguaje cotidiano o natural. Estos listados serán depurados mediante el traslado a un lenguaje normalizado, también denominado *lenguaje documental*. Una vez realizada la conversión, los términos pasarán a ser denominados *palabras clave* (BARROSO RUIZ, 1994: 128). Dentro de los lenguajes documentales se encuentran los tesauros.

La importancia de la terminología normalizada dentro de los procesos de repre-

sentación, procesamiento y transferencia del conocimiento se encuentra en continuo crecimiento, desempeñando un papel muy importante en el marco de la comunicación técnica y científica. Las nuevas tecnologías, la gestión digital de la información, la ingeniería del conocimiento y las ciencias relacionadas con la información han abierto nuevas perspectivas al ámbito terminológico, dotándolo de nuevas herramientas, métodos y criterios. Sin una terminología normalizada, es muy difícil el procesamiento de datos a través de las tecnologías de la transferencia de la información, convirtiéndose en una necesidad para llegar a una comunicación perfecta y una transmisión del conocimiento adecuado, ya que solo es posible si se entiende con exactitud el lenguaje en el que se efectúa dicha transferencia.

De ahí que especialistas de las distintas ramas del saber apuesten por un compromiso firme a la hora de utilizar los términos normalizados, comunicándose a través de «un mismo idioma», adoptando el lenguaje documental que les sea más apropiado. Este grupo de personas son las principales interesadas en que la comunicación, entre miembros de una misma especialidad, se pueda llevar a cabo con total efectividad. Lo mismo sucede cuando se cuenta con un lenguaje documental tipo tesoro para un sistema de información de un ámbito de conocimiento determinado. El personal técnico y especialista deberá contar con esta herramienta y utilizarla como base de la indización y de la sistematización de la información, aportando periódicamente, a la persona encargada de su mantenimiento, la nueva terminología para que forme parte del corpus del tesoro, implementando así una herramienta de trabajo en continua evolución.

A pesar de su importancia, la normalización terminológica está poco valorada dentro del desarrollo global de un sistema de información. Es un trabajo con es-

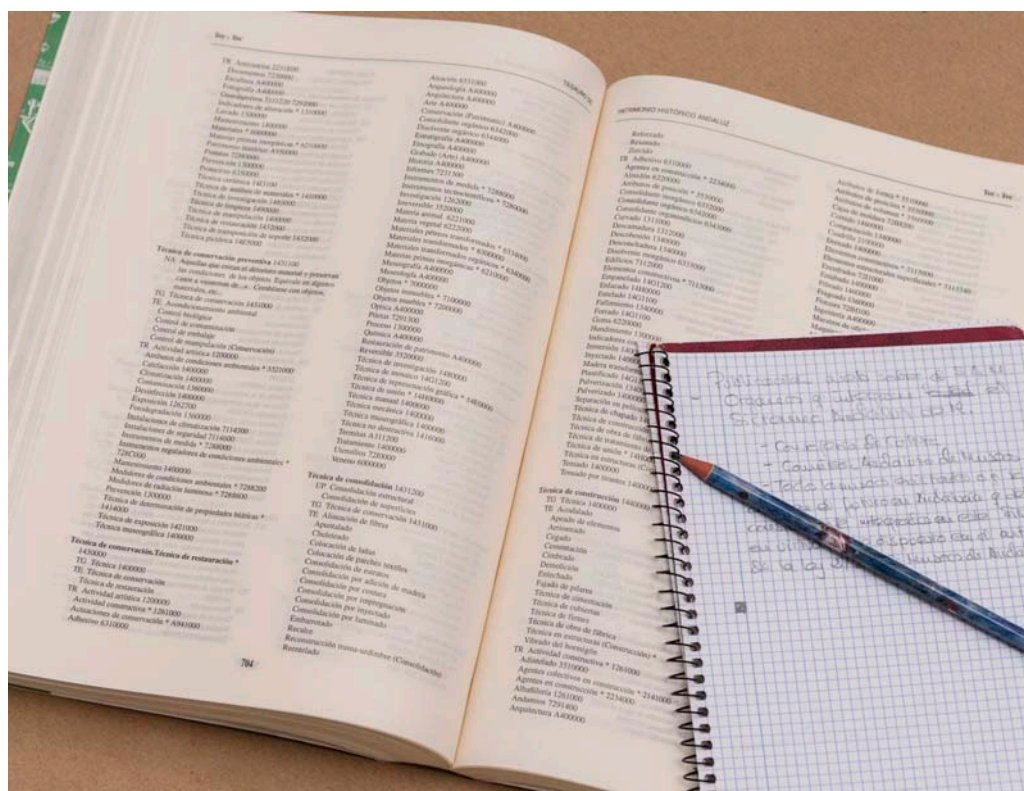
casa visibilidad, arduo, duro y lento, que siempre queda enmascarado por el conjunto visual de los sistemas de bases de datos. Sin embargo, sin este proceso no sería posible la representación y organización del conocimiento, provocando una indización con lenguaje libre, no controlado, que dificultaría las búsquedas produciendo ruidos o silencios.

Por otra parte, el desarrollo de las redes sociales ha permitido a las personas no solo aportar contenidos y productos web, sino que pueden también organizarlos aplicando las *folksonomías*, *tagging* o las etiquetas de lenguaje libre. Se trata de la utilización del lenguaje cotidiano por parte de las personas interesadas que acceden a estos espacios mediante una serie de términos que, en función de su mayor o menor utilización, pueden pasar o no a englobar el listado de términos normalizados, generados por el uso social (RODRÍGUEZ YUNTA, 2009: 121). Estas nubes de etiquetas podrían ser la base de un futuro tesoro, o una fuente para evaluar un lenguaje documental consolidado, observándose las carencias, necesidades de cambios y añadidos en su estructura, para adaptarse a la realidad social en el que se encuentra inmerso un sistema de información determinado.

ANÁLISIS DOCUMENTAL, INDIZACIÓN, TERMINOLOGÍA Y DESCRIPTORES

Como se ha comentado anteriormente, para llevar a cabo la transmisión de la información es necesario realizar de forma correcta su codificación sin ambigüedades, de forma clara y exhaustiva, identificando al objeto. Esto se consigue mediante el análisis del propio objeto estructurado en dos niveles:

- Nivel externo o formal: se efectúa mediante una descripción física que de-



Tesaurus. Foto: Isabel Dugo Cobacho

talla formalmente el objeto para su identificación y localización, independientemente de su significación científica y artística. Este análisis es similar al de los inventarios.

- Nivel interno o de contenido: se trata de profundizar en las características del objeto. Es asimilable a la catalogación, incluyendo su descripción (RUIZ PÉREZ, 1992: 32 y ss.).

La descripción desde distintas disciplinas científicas exigirá el uso de una terminología específica de cada una de ellas, relacionadas directa o tangencialmente con el patrimonio cultural. Tras este segundo paso se realizaría la indización.

La indización consiste en recopilar una serie de términos que se ajusten perfectamente a lo que representan los objetos. Es una de las operaciones centrales

del proceso documental e influye tanto en el almacenamiento como en la posterior búsqueda y recuperación de la información.

El resultado de la indización son los índices o listados de términos aplicados, mediante cajas de listas, en bases de datos o sistemas de información para satisfacer las necesidades de búsqueda de las personas interesadas. Las cajas de listas son el resultado del trabajo del personal especializado en cada una de las disciplinas relacionadas directa o indirectamente con el patrimonio cultural. En un principio, los listados se realizan en un lenguaje natural que, con posterioridad, hay que trasladar a un lenguaje normalizado, aunque si se cuenta con un thesaurus, se puede aplicar la terminología incluida en el corpus de ese lenguaje documental, por lo que se contará desde un principio

con una terminología normalizada (BARROSO RUIZ, 1994: 128).

La terminología puede estar formada por términos simples o compuestos, siendo denominados cada uno de ellos como *palabras clave*.

LENGUAJES DOCUMENTALES PARA NORMALIZAR Y RECUPERAR LA INFORMACIÓN

El lenguaje documental es el resultado de trasladar el lenguaje natural, libre o cotidiano, a un lenguaje normalizado y controlado, evitando todos los problemas derivados del uso de un lenguaje que cuenta con sinónimos, cuasisinónimos y polisemias. Si se utilizara un lenguaje libre, la búsqueda de información se vería enturbiada por resultados erróneos, entre los que se encontrarían los que realmente se quieren localizar, por lo que no se cumpliría la función de un sistema de información: localizar la información de forma rápida y eficaz.

Por el contrario, los lenguajes documentales surgen para evitar los problemas derivados de la utilización del lenguaje natural. Se trata de un conjunto de descriptores convencionales, controlados y normalizados, que se utilizan para representar y describir los objetos con la finalidad de facilitar la localización y recuperación de la información sin distorsiones.

Existen varios tipos de lenguajes documentales, cuya clasificación está recogida en diversos trabajos. De alguno de ellos ya se ha hablado, como son los lenguajes libres y lenguajes controlados; otros son los lenguajes precoordinados, y poscoordinados, los de estructura jerárquica, asociativa o combinatoria y los de estructura sintáctica.¹

Los tesauros se engloban dentro de un tipo de lenguaje documental intermedio que participa de la estructura jerárqui-

ca y asociativa. En cuanto a su clasificación, pueden ser facetadas, jerárquicas o mixtas.

¿QUÉ ES Y CÓMO SE CONSTRUYE UN TESAURO?

Definición y características

El término tesoro deriva de la palabra latina *thesaurus*, que significa «tesoro, depósito de riqueza y libro mágico». Se ha aplicado, sobre todo, en el ámbito anglosajón, a los diccionarios y otros libros que recopilan datos alfabéticamente. Pero, contrariamente a los diccionarios que ofrecen definiciones de palabras, los tesauros ofrecen relaciones existentes entre los términos que forman parte de su estructura.

En primer lugar, hay que determinar en qué consiste un tesoro y cuáles son las directrices que se utilizan para su construcción. Un tesoro no es un diccionario, por ello los descriptores no vienen definidos, sino determinados por su contexto (IAPH, 1998: 23). Es un lenguaje documental vivo y en continua evolución. La norma ISO 2788:1986, equivalente a la norma UNE 50106:1990, Directrices para el establecimiento y desarrollo de tesauros monolingües, establece una definición de este tipo de lenguaje documental: «Desde el punto de vista de su función, un tesoro es un instrumento de control de la terminología utilizada para trasponer a un lenguaje más estricto (lenguaje documental, lenguaje de información) el idioma natural empleado en los objetos por los indizadores y los usuarios. Desde el punto de vista de su estructura, el tesoro es un vocabulario controlado y dinámico, de términos que tienen entre ellos relaciones semánticas y genéricas, y que se aplica a un campo concreto del conocimiento».

En lo que respecta a las normas, hay que tener en cuenta, además, el contenido de la ISO 25964-1987, Directrices para la creación y desarrollo de tesauros multilingües, que reforma y amplía las normas sobre tesauros para su adaptación a los nuevos contextos de integración e interoperabilidad en el entorno digital.

Fruto del desarrollo de los tesauros y su aplicación a los sistemas de información fue la nueva norma UNE 25964-1: 2014, identificada con la norma ISO25964-1: 2011 y la UNE 25964-2: 2016, que se identifica con la norma ISO 25964-2: 2013. La norma 25964 se presenta bajo el título Información y documentación. Los tesauros y su interoperabilidad con otros vocabularios. Esta norma fue publicada en dos partes: la primera sobre tesauros para la recuperación de información, y la segunda sobre la interoperabilidad con otros vocabularios, publicada en 2016² (GARCÍA MARCO, 2016).

Una de las principales misiones del tesauro es eliminar la ambigüedad y sinonimia de los términos que componen su área temática, teniendo gran importancia su estructura jerárquica, de ahí que el tesauro surja como nexo de unión entre el objeto y el personal especializado, permitiendo pasar de un lenguaje natural a otro documental controlado que facilita el estudio y la catalogación del objeto, así como la rápida recuperación de la información.

El problema al que se enfrentan los tesauros tradicionales en la actualidad no es otro que el de la obsolescencia, debido, principalmente, a lo inamovible de la indización y de las relaciones internas establecidas. Las estructuras que se encuentran definidas en un tesauro hacen que solo se puedan recuperar los objetos que posean aquellos términos o relaciones con las que trabaja y que han sido definidas *a priori*.

Debido al desarrollo de las nuevas tecnologías, cada vez más avanzadas y con procesadores más rápidos, con mayor al-

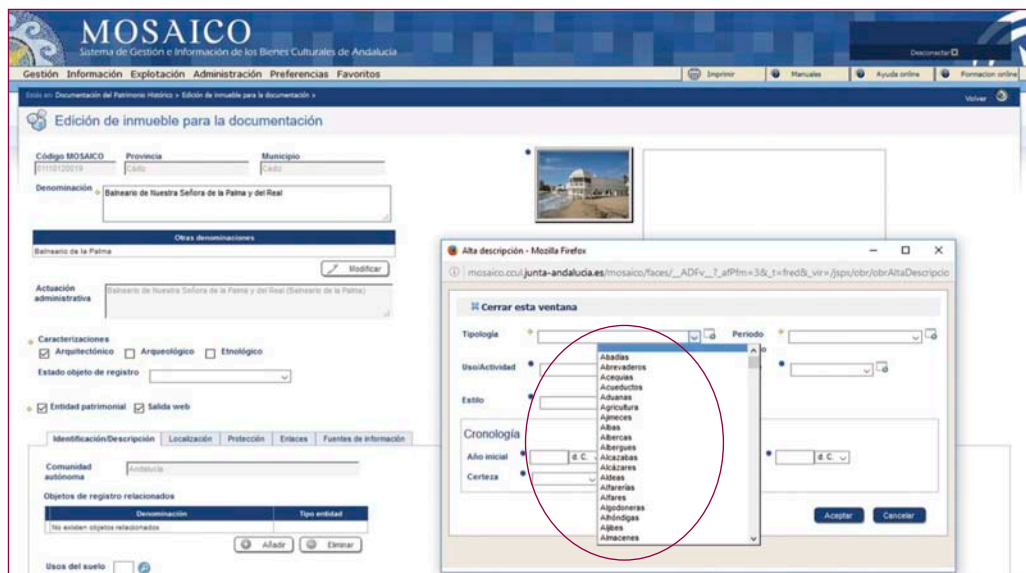
macenamiento de datos, más intuitivos y fáciles de manejar, los sistemas de información y los tesauros han debido adaptarse rápidamente a esta nueva realidad. Esto ha provocado que desde determinados sectores de la comunidad científica se soliciten cambios no solo estructurales en los tesauros, sino en el estudio y avance de las tecnologías enfocadas a la sistematización, localización y recuperación de la información (MARTÍNEZ GARCÍA, 2009; GARCÍA MARCO, 2016).

Construcción del tesauro

En la construcción de un tesauro, en primer lugar, se ha de tener en cuenta que el desarrollo de un lenguaje documental de este tipo es un proyecto que implica un gran esfuerzo tanto en recursos humanos como en el ámbito económico y temporal. Los principales trabajos que habría que desarrollar pueden estructurarse del siguiente modo:

1. Lo primero que se ha de llevar a cabo es una investigación exhaustiva para ver si existen experiencias similares desarrolladas por otras instituciones a nivel nacional e internacional. Esta búsqueda de información evitará, en la medida de lo posible, la repetición de esfuerzos innecesarios. La investigación se hará extensible a la bibliografía relacionada con los tesauros que se analicen y sean seleccionados para conseguir los objetivos planteados (MARTÍN PRADAS, 2007: 52).

Así, se estudiarán experiencias desarrolladas por instituciones relacionadas directa o tangencialmente con el patrimonio cultural, que podrán ofrecer pistas sobre la forma de enfocar el trabajo. Además, puede que se localice un tesauro que se adapte a las necesidades planteadas, parcial o totalmente, de ahí que en esta fase del



Cajas de lista. Tipologías de bienes inmuebles en el Sistema de Gestión e Información de los Bienes Culturales de Andalucía (Mosaico). Fuente: Consejería de Cultura. Junta de Andalucía

- trabajo puede finalizar porque se ha encontrado lo que se buscaba, o puede continuar con la realización de un tesoro propio.
2. Si se opta por realizar un tesoro propio, se han de volver a analizar en profundidad los tesoros y diccionarios afines al proyecto, aprovechando todos aquellos apartados que ya están incluidos en los tesoros consultados. Así se evitará, como se ha dicho antes, repetir un trabajo realizado con anterioridad que pueda resultar útil. De esta forma, no se parte de cero, sino de una experiencia consensuada, aprobada y probada en un sistema de información nacional o internacional. Para conseguir los objetivos habrá que centrarse en apartados concretos relacionados con el patrimonio cultural, desechando los campos que no interesen. Así se podrá constatar si hay campos importantes que han de formar parte del futuro tesoro y que no están recogidos en los consultados. En estos campos de nueva creación será donde se ha de aplicar el mayor esfuerzo del proyecto.

3. Para el desarrollo del tesoro es imprescindible formar un grupo de trabajo, encabezado por una persona que lo dirija y coordine. Este grupo de trabajo estará formado por especialistas de las distintas materias, disciplinas y ámbitos relacionados con el patrimonio cultural. Hay que tener en cuenta que a medida que se desarrollen los distintos campos conceptuales se va a observar la necesidad de ampliar el grupo e incluso contar con especialistas que lleven a cabo asesorías específicas.
- Mientras más centros, departamentos e instituciones relacionadas participan en el desarrollo y construcción del tesoro, más posibilidades de éxito tendrá y mayor aceptación y uso por la comunidad técnica del entorno.
4. Si se cuenta con un sistema de información, se partirá de la base de una serie de términos que han de convertirse en descriptores y no descriptores, es decir, sus sinónimos. Estos listados iniciales pueden ser, en principio, alfabéticos, pero es convenient-

te que cuando estén debidamente definidos y consensuados se les aplique la estructura jerárquica, es decir, de lo genérico a lo específico, e igualmente la estructura de equivalencia.

5. Una vez realizados los listados, divididos por campos o materias (algunos tesauros los denominan *facetas*), es necesario observar si son suficientes y es posible trabajar con ellos, implementándolos en función de las necesidades o, por el contrario, es necesario dar el siguiente paso para construir un verdadero tesoro.
6. Concluidos los trabajos de la fase anterior, los primeros listados quedarán perfectamente delimitados por los distintos grupos temáticos que cuentan con entidad propia, algunos de los cuales, debido a su contenido, pasarán a formar parte de la macroestructura del tesoro, integrándose como macrodescriptores de facetas. En estos primeros listados se pueden encontrar, por ejemplo, las facetas de estilos y períodos, materiales, técnicas, tipologías, objetos muebles, objetos inmuebles, etc. Estos apartados, que forman parte del cuerpo general del tesoro, deberán ir acompañados de una serie de listados asociados y auxiliares, también normalizados, que complementan la estructura principal del futuro tesoro.
7. A medida que se vaya desarrollando la recopilación de términos para las facetas antes mencionadas, se observará que hay términos que se repiten en las distintas facetas, siendo comunes a varias disciplinas. Esto va a desembocar en la elección definitiva del lenguaje documental tipo tesoro que se ha de desarrollar, permitiendo la integración de toda esta terminología que se repite. Un tesoro integrado, que trascienda los temáticos o disciplinares, propiciará un acercamiento al patrimonio desde diferentes

enfoques, posibilitando la descripción de un mismo objeto desde diferentes perspectivas: artística, arquitectónica, etnológica, arqueológica, etc. (Muñoz Cruz, 2007: 40).

Es necesario tener en cuenta las dificultades que pueden presentarse en el proceso de integración de la terminología procedente de las distintas ramas del conocimiento incluidas en el ámbito del patrimonio histórico dentro de un mismo lenguaje documental. En primer lugar, los intentos de integración pueden generar recelos entre especialistas de disciplinas diferentes. Por otro lado, es cierto que dicho esfuerzo por integrar implica la realización de concesiones, pero el resultado aportará más ventajas que inconvenientes. Estas concesiones vienen impuestas por la inclusión de determinados descriptores propios y compartidos con otras disciplinas en una determinada jerarquía, aunque siempre se puede llegar a acuerdos que den satisfacción al conjunto del equipo de trabajo (Martín Pradas, 2007: 53).

8. Una vez establecida la macroestructura y las facetas que la integran, el siguiente paso es adscribir los descriptores a cada uno de los campos establecidos, pasándolos de meros listados alfabéticos a listados jerárquicos, estableciendo la relación de equivalencia y, en la medida de lo posible, la relación asociativa. Por ello, cada especialista del grupo de trabajo ha de tener libertad para aportar términos relacionados de sus correspondientes disciplinas.
 - En primer lugar, deberá extraer los términos de la consulta de libros, diccionarios, enciclopedias e incluso del uso y necesidades que vea pertinentes en el desarrollo de su trabajo.

- Estos términos serán debidamente analizados en reuniones del grupo para ver su mejor adscripción dentro de las facetas que formaban parte de la macroestructura. También han de ser considerados como descriptores y seguidamente añadidos en el árbol escalonado de la jerarquía, aplicándose la metodología de lo genérico a lo específico, descendiendo niveles sin comprometer el significado de cada uno de ellos.
 - En aquellos casos en los que un descriptor presente resistencia de significado en una estructura, o por el contrario pueda ser ubicado en varios campos de la misma macrocategoría o en varias facetas, la solución más idónea será considerarlo polijerárquico (que aparece repetido en dos facetas o macrocategorías).
9. El siguiente paso es crear la estructura jerárquica, distribuyendo los términos en las grandes facetas que forman la macroestructura, siempre partiendo de lo genérico a lo específico. Para evitar la sinonimia se establecerá la relación de equivalencia y, paralelamente a todo el proceso, la estructura asociativa, relacionando los descriptores entre sí, en función de las necesidades: actividad-proceso-técnica-agentes-materiales-objetos, etc. (MARTÍN PRADAS, 2007: 53).
10. Para finalizar, habrá que comprobar que la aplicación de la terminología al sistema de información es fácil y se adapta perfectamente a sus necesidades. Del mismo modo, para el uso y manejo del tesoro hay que redactar unas normas para que cualquier persona interesada pueda comprender claramente la forma de utilizar este lenguaje documental.
- Durante las fases de construcción del tesoro se recomienda seguir la

metodología recogida en la Norma ISO 2788-1986E de Directrices para el establecimiento y desarrollo de tesauros monolingües, combinándose con una serie de actividades que la persona encargada de la coordinación del grupo irá aplicando. Algunas de estas actividades se llevarán a cabo antes, durante y después del desarrollo del tesoro y abarcan desde la definición de la amplitud temática del tesoro, hasta la revisión y modificación de la macroestructura, la creación y revisión de descriptores, la elaboración de notas de alcance, el establecimiento de relaciones de equivalencia, listados auxiliares y asociados, etc.

ESTRUCTURA, RELACIONES Y REPRESENTACIÓN DE UN TESORO

Una vez planteado qué es y cómo se construye un tesoro, es necesario profundizar en los elementos que lo conforman, sus relaciones internas, la polisemia y sinonimia y la representación de sus listados.

Elementos que componen el tesoro

El corpus del tesoro se compone de términos considerados descriptores o no descriptores.

- *Descriptores*: también llamados términos preferentes. Son aquellos términos que se validan en el tesoro y sirven para indizar el objeto. Son representaciones autorizadas de los conceptos sin posibilidad de polisemias o ambigüedades y pueden ser simples (una palabra) o compuestos (varias palabras).
- *No descriptores*: llamados también términos equivalentes o no preferentes. Son entradas no autorizadas del te-

sauro con un reenvío hacia los descriptores. Suelen ser sinónimos o cuasisinónimos de los descriptores y no pueden ser utilizados en la indización o para efectuar búsquedas. Los no descriptores son puntos de acceso no aceptados que reenvían a los descriptores autorizados (IAPH, 1998: 23-24).

Por último, respecto al género y número, sería conveniente revisar las normas, aunque hasta el momento en castellano siempre se ha usado el género masculino y el número singular para lo incontable y plural para lo contable.

Relaciones entre términos

Las relaciones que pueden mantenerse entre términos dentro de un tesaurus pueden ser:

1. Relación jerárquica: expresa la relación de superioridad y subordinación entre los conceptos dados por los descriptores, descendiendo desde lo genérico hasta lo más específico. Es la columna vertebral del tesaurus, ya que organiza su corpus. La notación empleada es TG (término genérico) y TE (término específico).
- TG: término genérico aplicado a aquellos con significado conceptual más amplio, que reagrupan términos específicos.
 - TE: término específico referido a aquellos cuyo concepto de significado es más concreto.

Ejemplo:

TG Períodos históricos
TE Prehistoria (TG)
TE Prehistoria reciente (TG)
TE Edad del Cobre



TG Períodos históricos. TE Prehistoria reciente (TG). TE Edad del Cobre. Ídolo antropomorfo. Jaén (España). Foto: Miguel Ángel Blanco de la Rubia

2. Relación de equivalencia: vincula entre sí términos que expresan el mismo concepto y se da entre descriptores y no descriptores. Dirigen desde el sinónimo, cuasisinónimo o antónimo hacia el descriptor autorizado. La reciprocidad se expresa a través de las convenciones USE y UP.

- USE: el término que lleva esta convención es el término preferente que se elige entre varios sinónimos.

Ejemplo:

Trabajadores USE Obreros
Asalariados USE Obreros
Operarios USE Obreros
Empleados USE Obreros

- UP: el término que lleva esta abreviatura es un sinónimo.

Ejemplo:

Obreros UP Trabajadores
UP Asalariados
UP Operarios
UP Empleados

3. Relación asociativa: se establece entre descriptores que no son equivalentes y que pueden ser relacionados jerárquicamente. Relaciona un descriptor con otro, ya sea por analogías entre descriptores o con el fin de precisar o ampliar un tema en el momento de la búsqueda documental. Esta relación es recíproca mediante la abreviatura TR.



TG Utensilios contenedores. TE Odres. TR Barriles. Bodegas Góngora. Villanueva del Ariscal, Sevilla (España). Foto: Fondo Gráfico IAPH (Juan Carlos Cazalla Montijano)

- TR: el término que porta este símbolo se asocia a otro término que no es su genérico, su específico o su sinónimo.

Ejemplo:

Odres TR Barriles
 TR Botas de vino
 TR Fermentación
 TR Maduración
 TR Vino

- NA: por último, existe otra abreviatura que se utiliza a la hora de construir tesauros: es el caso de la *nota de alcance*, NA, llamada también *nota de aplicación*. Se añade a algunos términos, o a la totalidad, con el fin de aclarar alguna particularidad, indicar combinaciones y ampliar o restringir el significado con el que es utilizado el descriptor. Sirve para seleccionar el descriptor más apropiado cuando se duda sobre la utilización de uno o varios descriptores candidatos (IAPH, 1998: 26).

Ejemplos:

Materiales.

NA: sustancias físicas naturales o transformadas por el ser humano.

Estudiantes.

NA: personas inscritas en una institución de enseñanza.

Exportación de bienes.

NA: temporal y permanente.

Bio.

NA: combínesse morfológicamente como prefijo con los descriptores correspondientes.

Respecto a las siglas, abreviaturas y acrónimos, si es posible, mejor no usarlos y, en el caso de que aparezcan, irán acompañados de

operadores de equivalencia que los autorizan o desautorizan mediante su correspondiente NA.

Ejemplo: bien de interés cultural
 USE BIC.

En la actualidad, las notas de alcance se presentan en algunos tesauros acompañadas de otras dos acepciones: por un lado, como nota de alcance de definición, y por otro, como nota de alcance histórica. Con esta diversificación se pretende ampliar el sentido con el que se usan los descriptores, muy recortados en la nota de alcance explicativa, ampliándolos con las posibilidades de la definición propia de cada término y, en su caso, una ampliación histórica en función de los conceptos o términos muy usados en otras épocas y en desuso en la actualidad. Con ello se consigue una visión más homogénea de cada término, obviando la gran carga de ambigüedad de la NA tradicional y ayudando al personal encargado de la catalogación en la elección de los descriptores.

Polisemias y sinonimias

El tesoro debe contener un vocabulario homogéneo y a la vez unívoco, que siempre cuenta con el mismo significado o interpretación. Para solventar aquellos casos en los que existan polisemias o sinonimias, se recomienda usar una serie de convenciones que permiten las normas (IAPH, 1988: 25):

1. Polisemia: homógrafos o términos polisémicos (a veces designados por el término general de *homónimos*). Misma ortografía, pero con significado diferente. Para eliminar los significados

múltiples pueden utilizarse diferentes procedimientos:

- Añadir calificadores o cualificadores, una palabra que acompaña al descriptor entre paréntesis. El calificador no forma parte del descriptor, no le acompañará en la base de datos o sistema de información, tan solo en el tesauro.

Ejemplo: avenidas (calles), manzanas (inmuebles), balancín (juguetes), ballestas (armas).



TG Estilos. TE Arte almohade. TR Actividad artística. Ermita de Cuatrovitás. Bollullos de la Mitación, Sevilla (España). Foto: Fondo Gráfico IAPH (Javier Romero García)

- Recurrir al plural y singular para evitar la polisemia:

Ejemplo: pulidores (agentes) en plural, y pulidor (utensilio) en singular.

- Utilizar el singular y el plural cuando la actividad y el lugar donde se desarrolla una actividad o trabajo se designen por el mismo nombre.

Ejemplo: carpintería y carpinterías.

- Incluir una pequeña aclaración en una nota de alcance (NA) de manera que pueda evitarse la polisemia y la ambigüedad del significado de una palabra.

Ejemplo: ceremoniales NA: libros en los que se trata sobre las ceremonias en los actos públicos de la Iglesia católica.

2. Sinonimia: se trata de palabras con igual significado y se controla a través de la relación de equivalencia antes mencionada. El operador USE reenvía de la/s entrada/s al término autorizado como se puede observar en el apartado dedicado a la relación de equivalencia.

Representación

La terminología del tesauro se puede representar de forma jerárquica, alfabética o con índice permutado. Cada una de ellas cuenta con el mismo número de descriptores y no descriptores.

1. La representación jerárquica es la representación del árbol jerárquico que contiene toda la terminología estructurada desde lo genérico a lo específico, descendiendo tantos niveles de

profundidad como el equipo de trabajo vea necesario establecer durante la elaboración del tesauro. Se representa mediante sangrados que indican el nivel de profundidad respecto a su término genérico, siguiendo clasificaciones establecidas o clasificaciones artificiales adoptadas por el grupo de trabajo.

Todos los términos pertenecientes a cada una de las jerarquías dependientes de una macrocategoría van acompañados de un código alfanumérico que ayudará a buscar el término en los listados alfasistemáticos y permutado en la publicación en papel (MARTÍN PRADAS, 2007: 54).

Ejemplo:

Estilo.

- Arcaísmo
- Art decó
- Arte abstracto
- Arte almohade
- Arte almorávide

2. En la representación alfabética o alfasistemática se recoge toda la terminología que forma parte del cuerpo del tesauro. En ella se presentan los términos genéricos (TG), términos específicos (TE), términos relacionados (TR), notas de alcance (NA) y sinónimos USE/UP.

Ejemplo:

Estilo.

NA Referido a estilos, artes, movimientos, escuelas y géneros artísticos.

TE Arcaísmo

TR Acabado

- Actividad artística
- Agentes colectivos
- Arte
- Atributos
- Ciudades

Estilo brutalista USE: Brutalismo

Estilo cubista USE: Cubismo

Estilo figurativo USE: Arte figurativo

Estilo funcionalista

TG Estilo

Estilo helenístico

TG Estilo

3. Al igual que los listados anteriores el índice permutado incluye todos los términos del tesauro, descriptores y no descriptores, alfabetizados por todas las palabras significativas que componen los términos, simples o compuestos. Es decir, un término aparecerá tantas veces como palabras contenga y en la posición alfabética que corresponda a cada una de ellas. Las entradas irían acompañadas del código alfanumérico (a la izquierda) y los sinónimos con reenvíos USE y UP.

Ejemplo:

Retoque abrupto.

Ábsides

Absidiolos

Absolutismo

Ensayo de absorción capilar

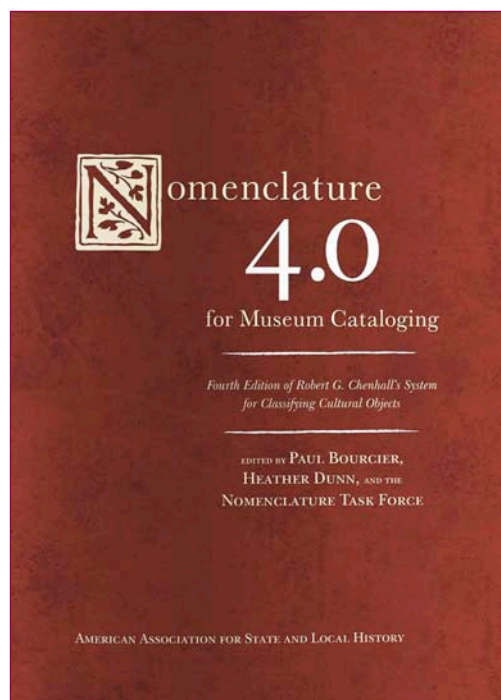
Ensayo de absorción libre

Reintegración por abstracción cromática

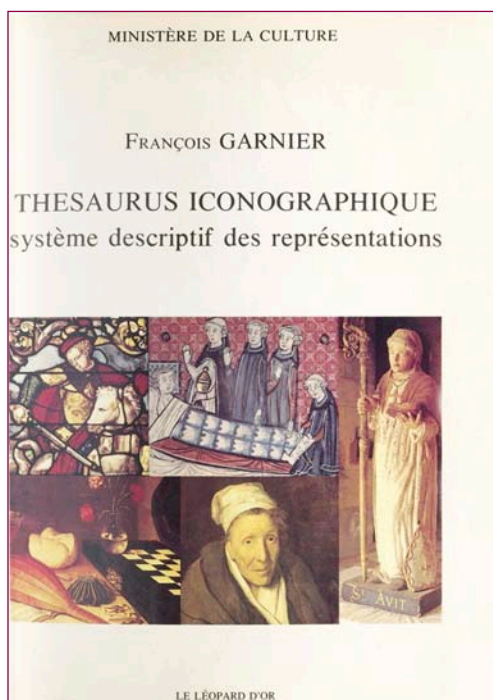
Arte abstracto

EXPERIENCIAS DE INTEGRACIÓN DE UN TESAURO EN UN SISTEMA DE INFORMACIÓN PATRIMONIAL

Aunque aparentemente son muchas las experiencias que se han dado en la integración de un tesauro en un sistema de información, realmente han sido pocas las que se han implantado en uno aplicado a la gestión de información de patrimonio cultural. Por el contrario, sí existen propuestas de tesauros que nacieron con esa vocación, aunque en muchos casos han sido traducidos y utiliza-



Cubierta de *Nomenclature for Museum Cataloging*. Paul Bourcier, Heather Dunn, and the Nomenclature task Force, 2015



Cubierta de *Thesaurus iconographique système descriptif des représentations*. París: Le Léopard d'or, 1984

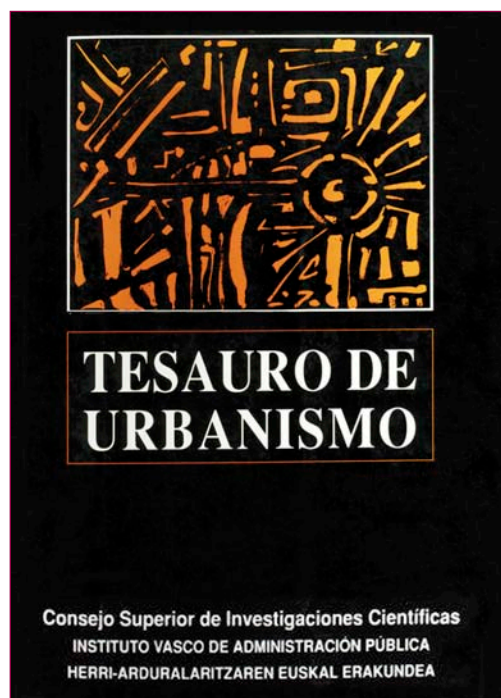
das algunas de sus partes, creando cajas de listas de términos aplicados a campos concretos de las fichas catalográficas de determinados sistemas de información o bases de datos.

Llegados a este punto, las instituciones no suelen tratar con claridad este tema, por lo que no se puede asegurar en qué medida han integrado la estructura, las jerarquías, determinados apartados e incluso han modificado algunas facetas adaptándolas a sus necesidades.

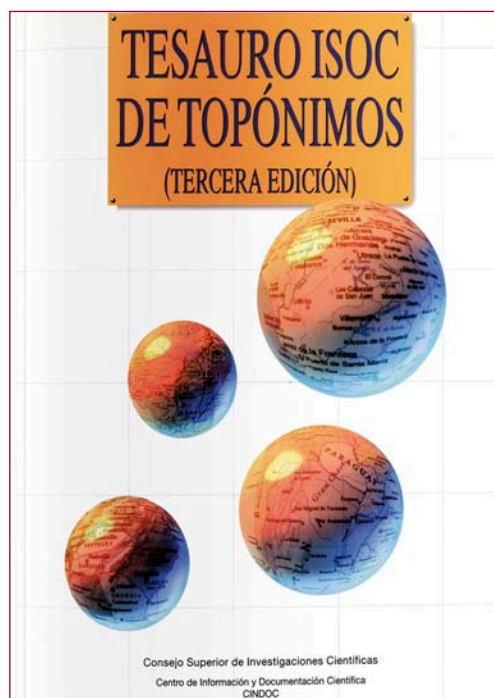
En este sentido, se intentará profundizar en qué medida los tesauros son aplicados verdaderamente a determinadas herramientas informáticas, ya que a veces se habla de aplicación de tesauros cuando se trata de listados que se van implementando en el desarrollo del propio trabajo, sin llegar a dar el paso siguiente para convertir estos listados en un lenguaje documental de este tipo. La finali-

dad de este apartado no es la de hacer un listado de todos los tesauros que se han realizado desde mediados del siglo pasado hasta la actualidad; por el contrario, el objetivo es el de dar a conocer algunas de las experiencias que han destacado por su innovación, efectividad, implantación, desarrollo y aplicación a un sistema de información.

Son muchos los tesauros que se han desarrollado para ser aplicados a determinados ámbitos relacionados con las humanidades, ciencias sociales, patrimonio cultural, arqueología, arte, historia, etc. Basta con acceder a cualquier buscador de Internet y establecer una búsqueda con la palabra «tesauro», para observar que aparece un elevado número de ellos. Muchos de estos tesauros han sido objeto de estudios en función de su temática, como el caso de los realizados por G. Mochón y Á. Sorli (2007), que se



Cubierta de *Tesouro de Urbanismo*. Madrid: Consejo Superior de Investigaciones Científicas, 1992



Cubierta de *Tesouro ISOC de Topónimos*. Madrid: Consejo Superior de Investigaciones Científicas, 2003

centran en los tesauros de humanidades y de ciencias sociales en Internet, donde se expone una idea general histórico-descriptiva y de contenido de los tesauros analizados.

Se encuentran tesauros como *Nomenclature for Museum Cataloging*;³ la clasificación iconográfica denominada Iconclass;⁴ el *Thesaurus iconographique système descriptif des représentations*;⁵ The British Museum Materials Thesaurus;⁶ English Heritage Monument type thesaurus;⁷ APT Australian Pictorial Thesaurus;⁸ Tesouro Architecture & Patrimoine;⁹ Tesouro Regional Patrimonial de Chile;¹⁰ Tesouro de la European Heritage Network (ABEJÓN PEÑA, 2016); *Tesouro de Urbanismo* (Isoc, 1992); *Tesouro ISOC de Topónimos* (CINDOC, 2003) y un largo etcétera.

Se exponen a continuación tres ejemplos de tesauros que fueron implantados con éxito en sistemas de información.

El AAT. Art & Architecture Thesaurus

Es uno de los más innovadores y punto de referencia obligado para el ámbito de la normalización terminológica a nivel mundial (PETERSEN, 1990). El AAT, junto con sus auxiliares ULAN (Lista Unificada de Nombres de Artistas) y TGN (Tesouro de Nombres Geográficos), constituye un lenguaje estructurado pensado para mejorar la gestión documental de la información relativa al ámbito de las artes y la arquitectura. La Fundación J. Paul Getty es la responsable de su elaboración y mantenimiento

Cuenta con 131 000 términos e informaciones complementarias en inglés con un alcance mundial y una cronología que abarca desde la Antigüedad hasta nuestros días. Cada uno de los descriptores se acompaña de un código numérico que lo

identifica, mostrando también el conjunto de relaciones asociativas, jerárquicas o de equivalencia que mantienen con otros términos, además de notas de alcance sobre su definición y uso.

Se estructura en facetas jerarquizadas que le aportan un esquema general de clasificación: conceptos asociados, atributos físicos, estilos y períodos, agentes, actividades, materiales y objetos. El ATT no solo 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 enfocado a museos, bibliotecas especializadas, archivos, catalogación de recursos visuales, proyectos bibliográficos relacionados con las artes y personal investigador en historia del arte. En su elaboración han colaborado varias instituciones.

Es este un caso de lenguaje documental cuya primera versión se realizó

en papel, aunque hoy día puede consultarse en línea.¹¹ Se caracteriza «por su alto grado de 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» (LAGUENS GARCÍA, 2006: 110-111).

El Tesauro de Arte & Arquitectura¹² es la traducción al español del AAT realizada por el Centro de Documentación de Bienes Patrimoniales, dependiente de la Dirección de Bibliotecas y Archivos de Chile, entre 2000 y 2003.

Tesauros y diccionarios de la Red Digital de los Museos estatales

Para la normalización de terminología y la unificación y estructuración del vocabulario técnico utilizado en la descripción y catalogación de bienes cul-

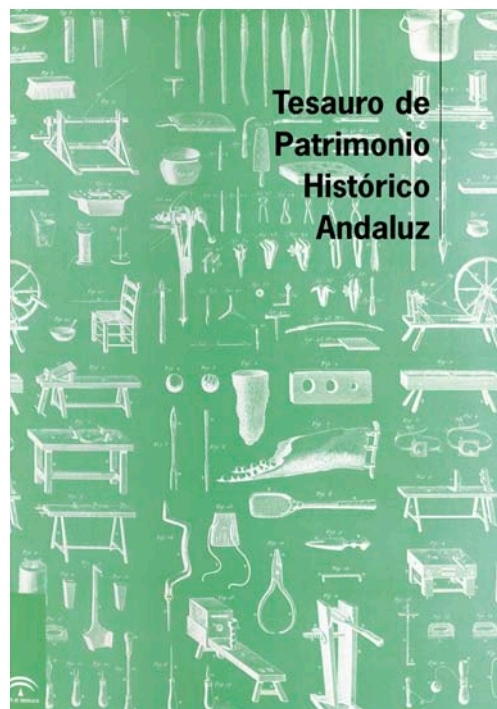
Red Digital de Colecciones de Museos Estatales. Tesauros del Patrimonio Cultural de España. Sede web del Ministerio de Educación, Cultura y Deporte

turales, se empezó a trabajar en 2001 en la elaboración de tesauros de patrimonio cultural, un proyecto en el que se sigue trabajando a día de hoy. Estos vocabularios se pueden clasificar en dos grupos: diccionarios especializados, que reúnen terminología específica de su correspondiente ámbito temático, y tesauros genéricos, que son aplicables a la catalogación de todo tipo de bienes culturales, muebles e inmuebles (ALONSO SÁEZ, 2012: 84 y ss.).

Los tesauros y diccionarios de la Red Digital de Colecciones de los Museos Estatales (CER.ES) responden al objetivo del Ministerio de Educación, Cultura y Deporte del Gobierno de España, de poner a disposición pública en Internet un catálogo colectivo de los bienes culturales de los museos de titularidad estatal y de los transferidos a las comunidades autónomas, a través del sistema integrado de documentación y gestión museográfica DOMUS. Proporciona también herramientas de normalización terminológica, mediante el control de los términos que representan a los conceptos y que se utilizan para describir, clasificar y gestionar los bienes culturales que custodian los museos mediante el uso de vocabularios controlados y tesauros (LAFUENTE URIEN, 2016: 181-182).

Los tesauros del patrimonio cultural realizados hasta el momento se han publicado en Internet y presentan más de 60 000 descriptores y 16 000 no descriptores pertenecientes a siete tesauros diferentes: materias, técnicas, mobiliario, numismática, cerámica, lugares geográficos y denominaciones de objetos culturales.

La nueva aplicación del Portal de Tesauros presenta la búsqueda de términos de forma alfabética y jerárquica, la visualización de imágenes asociadas a los términos o la recuperación de resultados en la Red Digital de Colecciones de Museos de España.¹³



Cubierta de *Tesauro de Patrimonio Histórico Andaluz*. Sevilla: [Consejería de Cultura] Instituto Andaluz del Patrimonio Histórico; Granada: Comares, 1998

Tesauro de Patrimonio Histórico Andaluz

El Tesauro de Patrimonio Histórico Andaluz (TPHA), elaborado por el Instituto Andaluz del Patrimonio Histórico, puede también destacarse por su vocación patrimonialista e integradora, siendo el primer tesauro que se desarrolló en España para ser aplicado a un sistema de información de patrimonio cultural (IAPH, 1998).

Los descriptores originarios fueron tomados de bases de datos sectoriales por disciplinas (arqueología, arquitectura, etnología y arte), de las que partieron una serie de listados de términos jerárquicos y alfabéticos con relaciones de equivalencia.

La publicación del TPHA supuso un antes y un después en el tratamiento de la información y en la eficacia de su poste-

rior recuperación. Fue a partir de este momento cuando se inició la aplicación del nuevo lenguaje documental al Sistema de Información del Patrimonio Histórico Andaluz (SIPHA) (Muñoz Cruz, 2007: 40).

Es un tesoro estructurado en facetas donde se clasifican las disciplinas relacionadas con el patrimonio cultural, acompañados de listados asociados y auxiliares que completan y complementan el cuerpo del tesoro. Tiene como objetivo, por tanto, ser de utilidad a todas aquellas personas e instituciones relacionadas con el conocimiento y la sistematización de la información del patrimonio cultural en lengua española, constituyendo un instrumento capaz de adaptarse a las necesidades planteadas desde los distintos ámbitos patrimoniales y contextos territoriales. Alcanza un total de 16 288 descriptores, 13 948 relaciones entre descriptores, 1583 sinónimos, 1761 notas de alcance y 51 facetas.

Hay que destacar también que el TPHA ha sido utilizado en el desarrollo de varios proyectos, consiguiendo varios de los objetivos inicialmente marcados (Martín Pradas, 2007: 60):

- Comprobar su validación como herramienta documental en campos específicos de investigación relacionados con el patrimonio cultural.
- Obtener una mayor difusión en instituciones nacionales e internacionales.
- Traducir parte de la terminología utilizada a otros idiomas.
- Implementar la terminología con nuevos términos.

El TPHA se encuentra accesible para su consulta en línea¹⁴ y, en la actualidad, se está construyendo una nueva aplicación para su gestión y difusión con el objetivo de favorecer su reutilización y la interoperabilidad de los datos (datos abiertos). En esta nueva aplicación se integrará también su traducción al inglés.

NOTAS

1. Para profundizar en los tipos de lenguajes documentales clasificados según varios puntos de vista puede consultarse: MARTÍN GAVILÁN, 2009.
2. Norma UNE-ISO 25964-2:2016 Interoperabilidad con otros vocabularios <<http://www.aenor.es/aenor/normas/normas/fichanorma.asp?tipo=N&codigo=N0057671#.WodmVtLyIM8>> [Consulta: 26/04/2017].
3. Nomenclature 4.0 for Museum Cataloging <<http://resource.aaslh.org/view/nomenclature-4-0-for-museum-cataloging/>> [Consulta: 26/04/2017].
4. Iconclass <<http://www.iconclass.nl>> [Consulta: 26/04/2017].
5. Thesaurus iconographique, système descriptif des représentations de François Garnier <<http://www.culture.gouv.fr/documentation/joconde/fr/partenaires/AIDEMUSEES/thesaurus-garnier/thesaurus-pres.htm>> [Consulta: 26/04/2017].
6. British Museum Materials Thesaurus <<http://www.vocabularyserver.com/materials/index.php>> [Consulta: 26/04/2017].
7. English Heritage Monument Type Thesaurus <http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=1> [Consulta: 26/04/2017].
8. APT Australian Pictorial Thesaurus <<http://www.vocabularyserver.com/apt/>> [Consulta: 26/04/2017].
9. Thésaurus de la désignation. Architecture & Patrimoine <<http://www.culture.gouv.fr/culture/inventai/patrimoine/>> [Consulta: 26/04/2017].
10. Tesoro Regional Patrimonial de Chile <<http://www.tesaurorregional.cl/trp/publico/buscar.htm>> [Consulta: 26/04/2017].
11. Getty Vocabularies <<http://www.getty.edu/research/tools/vocabularies/index.html>> [Consulta: 26/04/2017].
12. El Tesoro de Arte & Arquitectura <www.aatespanol.cl> [Consulta: 26/04/2017].
13. Tesoros del Patrimonio Cultural de España. Red Digital de Colecciones de Museos de España <<http://tesoros.mecd.es/tesoros>> [Consulta: 26/04/2017].
14. El Tesoro de Patrimonio Histórico Andaluz <<http://www.iaph.es/web/canales/conoce-el-patrimonio/tesauro-pha/pagina2.html>> [Consulta: 26/04/2017].

BIBLIOGRAFÍA

- ABEJÓN PEÑA, T. (2016) Dimensión política y cultural en la armonización terminológica multilingüe: la experiencia del tesoro para la red de

- información del Patrimonio Europeo (HERE-IN). En *Actas del curso El Lenguaje sobre el patrimonio. Estándares documentales para la descripción y gestión de colecciones*. Madrid: Ministerio de Educación, Cultura y Deporte, 2016
- ALONSO SÁEZ, R. (2012) Tesoros de Patrimonio cultural: conocimiento en línea al servicio del ciudadano. SIAM. *Series Iberoamericanas de Museología* [en línea], Año 3, vol. 1, 2012, pp. 83-89. <<http://hdl.handle.net/10486/11388>> [Consulta: 26/04/2017]
- BARROSO RUIZ, S. (1994) La Normalización terminológica en los museos. El Tesoro. *Revista General de Información y Documentación* [en línea], vol. 4, n.º 2, 1994, pp. 121-160. <<http://revistas.ucm.es/index.php/RGID/article/view/RGI-D9494220121A/11531>> [Consulta: 26/04/2017]
- CINDOC (2003) *Tesoro ISOC de Topónimos 2003*. Madrid: Centro de Información y Documentación Científica. CSIC, 2003
- GARCÍA MARCO, J. (2016) Normas y estándares para la elaboración de tesoros de patrimonio cultural. En *Actas del curso El Lenguaje sobre el patrimonio. Estándares documentales para la descripción y gestión de colecciones*. Madrid: Ministerio de Educación, Cultura y Deporte, 2016, pp. 29-46
- IAPH (1998) *Tesoro de Patrimonio Histórico Andaluz*. Sevilla: Instituto Andaluz del Patrimonio Histórico. Junta de Andalucía, 1998
- ISOC (1992) *Tesoro ISOC de Urbanismo*. Madrid: Instituto de Información y Documentación en Ciencias Sociales y Humanidades. CSIC, 1992
- LAFUENTE URIEN, A. (2016) PARES hacia la web semántica: autoridades y tesoros en los archivos estatales. En *Actas del curso El Lenguaje sobre el patrimonio. Estándares documentales para la descripción y gestión de colecciones*. Madrid: Ministerio de Educación, Cultura y Deporte, 2016, pp. 181-194
- LAGUENS GARCÍA, J.L. (2006) Tesoros y lenguajes controlados en Internet. *Anales de Documentación* [en línea], n.º 9, 2006, pp. 105-121. <<http://revistas.um.es/analesdoc/article/viewFile/1391/1441>> [Consulta: 26/04/2001]
- MARTÍN GAVILÁN, C. (2009) Lenguajes documentales, principales tipos de clasificación, encabezamientos de materia, descriptores y tesoros. Temas de biblioteconomía [en línea]. 2009. <<http://eprints.rclis.org/14817/1/lendoc.pdf>> [Consulta: 26/04/2001]
- MARTÍN PRADAS, A. (2007) El Tesoro de Patrimonio Histórico Andaluz como lenguaje integrador. *El Sistema de Información del Patrimonio Histórico de Andalucía (SIPHA)*. Sevilla: Instituto Andaluz del Patrimonio Histórico, 2007, pp. 50-61 (PH cuadernos, n.º 20)
- MARTÍNEZ GARCÍA, S. (2009) La representación y organización de la información a través de los tesoros. La interdisciplinariedad como nuevo paradigma: Retos para la Documentación y la Biblioteconomía. *Contribuciones a las Ciencias Sociales* [en línea], n.º 6, 2009. <<http://www.eumed.net/rev/cccss/06/smg.htm>> [Consulta: 26/04/2017]
- MOCHÓN BEZARES, G.; SORLI ROJO, A. (2007) Tesoros de Ciencias Sociales en Internet. *Revista Española de Documentación Científica* [en línea], vol. 30, n.º 3, 2007, pp. 395-419. <<http://redc.revistas.csic.es/index.php/redc/article/view/392/404>> [Consulta: 26/04/2017]
- MUÑOZ CRUZ, V. (2007) Un modelo de integración y normalización. *El Sistema de Información del Patrimonio Histórico de Andalucía (SIPHA)*. Sevilla: Instituto Andaluz del Patrimonio Histórico, 2007, pp. 30-49 (PH cuadernos, n.º 20)
- PETERSEN, T. (dir.) (1990) *AAT, The Art and Architecture Thesaurus*. Paul Getty Foundation. New York: Oxford University Press, 1990
- RODRÍGUEZ YUNTA, L. (2009) Etiquetado libre frente a lenguajes documentales. Aportaciones en el ámbito de Biblioteconomía y Documentación. En *IX Congreso ISKO-España. Nuevas perspectivas para la difusión y organización del conocimiento. Valencia 11-13 de marzo de 2009*. Valencia: Universidad Politécnica de Valencia, 2009, pp. 832-845
- RUIZ PÉREZ, R. (1992) *El análisis documental: bases terminológicas, conceptualización y estructura operativa*. Granada: Universidad de Granada, 1992