
Una base de datos de saber terminológico para las artes espaciales

A terminological knowledge database for the fine arts

Claudio MOLINA SALINAS

Instituto de Investigaciones Estéticas de la Universidad Nacional Autónoma de México, Circuito Mario de la Cueva S/N, Ciudad Universitaria, 04510, Ciudad de México, México. claudio.molina.salinas@comunidad.unam.mx

Resumen

En este artículo se presenta el desarrollo conceptual de una base de datos de terminología de las artes espaciales que recogerá usos del español nacional mexicano. Esta brindará certeza del uso del vocabulario de especialistas del dominio, permitirá reconocer las terminologías comunes y diferenciadas usadas por especialistas mexicanos, apoyará las prácticas descriptivas y de normalización de datos que se generan en el sector cultural de México, y propiciará la normalización de las bases de datos existentes en las instituciones de la memoria de México. Por último, cabe señalar que, para el diseño conceptual de la base de datos, se han considerado principios propios de la lingüística y sus subdisciplinas aplicadas, así como de las ciencias de la información; además, se espera que esta iniciativa sea una primera base documental que funja como control terminológico y que, en un futuro, se convierta en un tesoro nacional para el dominio.

Palabras clave: Bases de datos terminológicas. Artes visuales. Artes espaciales. Patrimonio cultural. Control terminológico. Control de vocabulario. Patrimonio terminológico. México.

1. Introducción

La propuesta conceptual que se desarrolla en este artículo ha considerado aspectos teóricos y metodológicos propios de la terminología, terminografía y terminótica que se han aplicado a la construcción de una base de datos de saber terminológico; misma que servirá como un incipiente control del vocabulario de las artes espaciales en México (1).

Aunque el objetivo final de este artículo es la descripción conceptual de la base de datos, previo a la presentación de esta se discute la importancia de considerar los nuevos entornos digitales como un medio de control de la terminología del arte, las razones por las que se considera necesario contar con una base de datos de saber terminológico para México, y se ejemplifican estas problematizando a partir de ejemplos puntuales del *Tesoro de Arte & Arquitectura* y los *Tesoros-Diccionarios del patrimonio cultural de España*. Además, se ofrecen definiciones de conceptos

Abstract

This article presents the conceptual development of a database of fine arts terminology that will gather uses of national Mexican Spanish. This database will provide certainty of the use of the vocabulary of specialists in the domain, allow for the identification of the common and differentiated terminologies used by Mexican specialists, support the descriptive and data normalization practices that originate in the cultural sector of Mexico, and favour the normalization of databases present in Mexican museums, archives, libraries and galleries. Lastly, it is worth mentioning that for the conceptual design of the database, principles related to linguistics, its applied sub-disciplines, and information sciences have been considered; this initiative is expected to be a first documentary basis which acts as terminological control and that, in the future, it will become a national thesaurus for the domain.

Keywords: Terminological databases. Visual arts. Fine arts. Cultural heritage. Vocabulary control. Heritage terminology. Mexico..

como tesoro, sistema conceptual y sistema simple de organización del conocimiento, para dar un marco teórico general que ayude a entender los alcances y pormenores de la propuesta.

Hacia el final del documento se explica lo que se entiende por base de datos de saber terminológico, se abordan aspectos específicos sobre la compilación y estructuración del modelo, y se proyectan futuras líneas de trabajo.

2. Entornos digitales, terminología del arte y vocabularios controlados

El uso creciente del internet y sus alcances implican un cambio notable en la apertura de datos y la masificación del conocimiento (Rodríguez Ortega, 2016); este cambio epistémico responde a las nuevas dinámicas sociales en las que la Web 2.0 se ha convertido en un espacio social, de esparcimiento y de encuentro humano (Combi, 2016).

Desde una perspectiva antropológica, se consideran *digital natives* a quienes interactúan mayormente en estos entornos; a saber, individuos nacidos después del año 1980 que, según sus prácticas comunicativas, sociales y de esparcimiento, privilegian la inmediatez, el acceso y la fácil obtención y transferencia de información (Combi, 2016).

En México, según la *Encuesta Intercensal 2015: principales resultados* (INEGI, 2015) del Instituto Nacional de Estadística y Geografía (INEGI), en el año 2015 los *digital natives* superaban el 53% de la población total. Considerando esto, resulta natural asumir que, a corto plazo y para el caso mexicano, el uso e implementación de herramientas informáticas y la difusión del conocimiento en ambientes digitales serán inminentes, además de que la apertura de datos, la masificación del conocimiento y su difusión en línea serán una demanda de la gran mayoría de la población.

En el terreno de la gestión del patrimonio cultural, área que se vincula con el desarrollo de esta propuesta, experimentamos un paradigma novedoso en cuanto a la descripción, clasificación y estructuración de datos para su uso, reutilización y aprovechamiento efectivos en la gestión de colecciones digitales y acervos patrimoniales de instituciones de la memoria (Rodríguez Ortega, 2016).

En esta área del conocimiento, se tiene por cierto que hay cuatro condiciones fundamentales que requieren las bases de datos de bienes culturales: estructuras de datos (*data fields & structure*), contenidos descritos en estas estructuras (*data content & values*), etiquetas de marcación que permitan la clasificación y fácil recuperación de la información (*data format*), y un protocolo específico que facilite compartir esta información (*data exchange*). Lo anterior propicia y facilita la explotación automatizada y la reutilización del conocimiento de este dominio del saber (Elings & Waibel, 2007).

Los contenidos descritos (*data content & values*), en la mayoría de los casos, remiten a vocabularios controlados que restringen, desambiguan, normalizan, jerarquizan y relacionan terminologías (Harping, 2010b; Barité Roqueta, 2014) y, a su vez, permiten homologar la descripción y catalogación del patrimonio, y la obtención del conocimiento sobre este (Harping, 2010b).

Según Harping (2010a), los tesauros, en particular los del Proyecto de Tesauros del Instituto Getty, tienen como función principal catalogar, recuperar información y fungir como una herramienta de búsqueda para catalogadores e indexadores; esto se consigue mediante la estructuración de ellos con base en la norma ISO 25964,

correspondiente a la normatividad para la construcción de tesauros y su interoperabilidad con otros instrumentos de control terminológico.

The Art & Architecture Thesaurus (A&AT) y su versión en español, *The Getty Thesaurus of Geographic Names* (TGN), *The Union List of Artist Names* (ULAN) y *The Cultural Objects Name Authority* (CONA) son tesauros ampliamente usados por comunidades de catalogación; sin embargo, estas no son las únicas iniciativas actuales y de amplio alcance sobre la normalización de vocabularios referidos al patrimonio, ya que, para el caso específico de la lengua española, también existe el proyecto de los *Tesauros-Diccionarios del Patrimonio Cultural de España* (TDPCE), solo por mencionar un ejemplo.

Centrándonos en este último caso, es claro que el proyecto manifiesta la importancia de la normalización de vocabularios y su sistematización en tesauros con miras a la descripción y gestión de colecciones del patrimonio cultural (2). El resultado de esta iniciativa se traduce en que actualmente se cuenta con tesauros generales de materias, denominaciones, lugares, técnicas, así como tesauros específicos de cerámica, numismática y mobiliario (Ministerio de Cultura y Deporte, n.d.).

Desafortunadamente, se ha identificado que la implementación de tesauros como el TA&A y los TDPCE, en el uso del contexto de México, trae consigo problemáticas asociadas a la ambigüedad de las definiciones y su precisión, a la cobertura terminológica de las obras de consulta, la localización de las definiciones y adecuación a los usos locales, la organización de los términos en las facetas, y la armonización con los sistemas de catalogación actuales. Por ello, a continuación, se ilustrarán las razones por las cuales se considera necesario desarrollar una iniciativa propia para el contexto mexicano.

3. La necesidad de un tesoro de las artes espaciales para México

En esta propuesta de trabajo se plantea la idea de desarrollar, en una primera etapa, una base de datos de saber terminológico para las artes espaciales en México, que se convertirá, a corto plazo, en una herramienta de documentación terminológica del área y un control de autoridad y, a mediano y largo plazo, será la base documental terciaria para la construcción de un tesoro propio para México.

Un proyecto de semejante envergadura sería cuestionable al pensar que ya existen, cuando menos, otras dos iniciativas de control terminológico en español, mismas que ya se han referido en el apartado anterior. Pero, contrariamente a lo

que se podría pensar, la abundancia de diccionarios técnicos, obras de consulta y controles terminológicos no conducen a la normalización del vocabulario de un dominio del saber humano (Muñiz Castro, 2004); si este fuera el caso, el problema de la variabilidad terminológica sería una asignatura superada desde hace muchos años. Para ser más claro al respecto, me permito referir al término *bacín*, documentado en el TA&A. Como se puede apreciar en la “Nota de aplicación” del ejemplo citado, un bacín o bacines son (Centro de Documentación de Bienes Patrimoniales, n.d.):

Vasijas redondas y abiertas anteriormente usadas para contener líquidos o sólidos, hoy en día principalmente para contener líquidos. En Inglaterra en el siglo XVII el término se refería a una fuente poco profunda, mientras que ahora allí [SIC] se aplica a un bol profundo. En los Estados Unidos el término siempre ha significado una fuente profunda o un bol poco profundo.

Como se puede ver, la “Nota de aplicación” se aproxima de forma incorrecta al significado del término *bacín* y, por lo tanto, la explicación resulta ambigua, ya que la definición remite a tres significados distintos: (a) la vasija que puede contener sólidos y líquidos, (b) un tipo de fuente poco profunda y (c) un bol.

Este ejemplo corrobora la afirmación de que la obra de consulta no es suficiente para normalizar el concepto; esto, desde dos puntos de vista: primero, se asocian tres contenidos a un solo término, lo que se define como polisemia, entendida como “el fenómeno que consiste en que una palabra tenga [...] más de un significado (lo cual se muestra claramente en los diccionarios)” (Lara Ramos, 2006: 106), y que, para el control de vocabularios, es totalmente indeseado, pues propicia ambigüedad conceptual (Cabré i Castellví, 1998); y segundo, el alcance de la definición.

Para defender este último punto de vista, me permito llamar la atención a la definición glosada en la letra (a) y hacer la pregunta siguiente para probar la efectividad de la definición: ¿Todas las “vasijas redondas y abiertas anteriormente usadas para contener líquidos o sólidos, hoy en día principalmente para contener líquidos”, son “bacines”?

Un “no” por respuesta es suficiente para afirmar que la definición presentada es tan general que propicia la ambigüedad conceptual debido a que, al estar formulada como pregunta, se podrían asociar a ella otros términos como “cuenco”, “escudilla”, “pote”, “tazón”, entre otros.

El procedimiento descrito antes consiste en convertir una definición en una pregunta y, a partir de ella, recuperar un término o palabra. Esta prueba resulta muy común en lexicografía general y especializada, y permite emitir un juicio sobre la

pertinencia de una definición, ya que referir a otro vocablo significará que la definición es imprecisa (Lara Ramos, 2006, p. 37-52).

Esto aparte, en el TA&A no se encuentran todos los términos referidos al patrimonio cultural de México; por ejemplo, el término *mojigango* (3) no se registra. Esto es un indicio de que el tesoro no tiene un alcance normalizador para México y, en consecuencia, carece de cobertura de los términos de uso local.

Por otra parte, es frecuente encontrar términos que no están definidos como se entienden en una comunidad de práctica, o como un especialista lo requeriría para la catalogación de un objeto; por ejemplo, “caseína” se explica así (Centro de Documentación de Bienes Patrimoniales, n.d.):

Una proteínas [SIC] de la leche natural que contiene fósforo compuesto por aminoácidos cuajada de caseína [SIC] en forma natural como leche agria, pero también es precipitado por la adición de ácido clorhídrico diluido en leche descremada caliente. La cuajada se recoge, se lava y se seca para formar un polvo blanco o amarillento. La caseína seca es insoluble en agua y alcohol, pero es soluble en los carbonatos y otras soluciones alcalinas. Para el uso, la caseína se remoja durante la noche en una solución con un álcali débil (carbonato de amonio, bórax o cal) para formar una solución clara, viscosa [...].

En esta definición se ofrece un significado que remite al concepto; sin embargo, la explicación omite enunciar la trascendencia y aplicaciones de la “caseína” como un material usado en la ebanistería, como pegamento y para aumentar la durabilidad de barnices y, en la pintura, como una imprimación para bases de madera y de otros materiales, según lo define Doener (2005).

Por último, me permito señalar que las facetas del TA&A podrían resultar poco claras para los hispanohablantes, ya que la versión en español es una traducción del inglés y, por eso, el ordenamiento y jerarquización de la terminología que está detrás del facetado del tesoro reflejan una concepción e interpretación anglocéntrica del conocimiento (4). Ya que recurrir a un sistema de facetas para organizar el vocabulario de un área del conocimiento es una alternativa correcta para la recuperación de términos y la sistematización del vocabulario de un dominio del saber, pero suponer que los conceptos emanan de una realidad completamente estructurada, común a todas las lenguas, podría ser impreciso. En virtud de lo anterior, se plantea que la organización de los términos de una disciplina no debería sistematizarse a partir de un conjunto de conceptos delimitados en un mundo objetivo dado de antemano, ya que las lenguas se manifiestan en versiones verbales intuitivas e independientes del mundo objetivo (Coseriu, 1985; Fischer, 2004).

Por otro lado, una alternativa al uso del TA&A sería adoptar como control terminológico cualesquiera de los TDPCE; empero, una de las primeras limitaciones que encuentro en este proyecto es la disponibilidad de tesauros, ya que en la actualidad solo se cuenta con un *Diccionario de materias*, un *Diccionario de denominaciones de bienes culturales*, un *Diccionario geográfico* y un *Diccionario de técnicas*, considerados como parte de un conjunto llamado “tesauros generales” y, por otro lado, tres diccionarios o tesauros específicos: el *Diccionario de cerámica*, el *Diccionario de numismática* y el *Diccionario de mobiliario*.

También conviene considerar que la difusión geográfica de nuestra lengua, en 21 países, diversifica los usos de los términos y su significación (Muñiz Castro, 2004), por lo que adoptar un conjunto de tesauros desarrollados en España implica una notable falta de registros de términos de uso local, giros o matices de significado.

En cuanto a la posibilidad de adoptar como control terminológico otras obras de consulta, se ha identificado que algunos de los trabajos en línea retoman significados de diccionarios de la lengua general, lo que supone un error implícito en la descripción de los conceptos y, por lo tanto, se daría cuenta parcial del significado especializado de los términos (Molina Salinas, 2017).

Adicionalmente, hay que considerar que los tesauros y las obras de consulta en papel no están exentos de todas las observaciones que se han enunciado y, en la mayoría de los casos, resultan incompatibles con los sistemas de catalogación actuales, ya que no han sido pensados para ofrecer contenidos estandarizados, como un número único de identificación para cada término, una URL única que permita tener acceso al término en el vocabulario que lo describe, entre otros (Molina Salinas, 2017).

A partir de estos planteamientos, se ha considerado construir un tesoro normalizado, localizado y organizado mediante un facetado que refleje la concepción e interpretación de la teoría y el ejercicio de la catalogación en México, que defina unívocamente a los términos y que se sistematice a partir de los conceptos e interdisciplina de la teoría de la historia del arte y museografía, la lexicología y terminología, y de las ciencias de la información.

4. Tesauros y sistemas conceptuales en internet: documentando terminologías

El desarrollo del campo interdisciplinar encargado del estudio y aplicación de la tecnología al procesamiento de la información ha dado pie a la interacción de disciplinas humanas con aplicadas y metodologías de las ciencias. En este sentido,

la terminología y terminografía han incorporado sus metodologías en contextos de disciplinas como la bibliotecología, historia del arte y las tecnologías de la información. Considerando esta afirmación, en este apartado se presentan algunos conceptos de la bibliotecología estrechamente relacionados con nociones propias de la terminología y las tecnologías de la información que resultan fundamentales para la catalogación del patrimonio cultural, a saber: la consolidación de tesauros y sistemas conceptuales normalizados que funcionen en un entorno web.

Aunque los tesauros nacen en el seno de la lexicografía inglesa, en 1852, cuando Peter Mark Roget publicara el *Roget's Thesaurus of English Words and Phrases*, no es sino hasta 1957 que en el campo de la bibliotecología y la documentación se usan como herramientas para analizar la problemática de traducir conceptos y evidenciar las relaciones sistemáticas que establecen entre ellos (Arano, 2005). A partir de la década de los sesenta del siglo pasado, los bibliotecólogos y científicos de la información adoptarían a los tesauros como (Arano, 2005):

un tipo de lenguaje documental que representa la estructuración conceptual de un determinado campo del conocimiento [que] proporciona una organización semántica principalmente a través de la explicitación de las relaciones establecidas entre dichos conceptos.

Hoy en día, los tesauros constituyen un lenguaje documental que permite recuperar información sobre contenidos temáticos, un lenguaje de indexación que agrupa a objetos por temas y, lo que es más relevante para esta propuesta, funge como un lenguaje controlado que explicita una terminología normalizada, la estructura conceptual de un campo temático y evidencia las relaciones entre los conceptos de un dominio (Nau-mis Peña, 2007).

Por otra parte, los pasos o secuencias de trabajo para la creación de una obra de consulta especializada, en este caso un tesoro, son: (i) la definición del trabajo, (ii) la preparación del trabajo, (iii) el diseño, la construcción y la explotación de corpus, (iv) la gestión terminológica, (v) la revisión y supervisión del trabajo y (vi) la edición (Alpizar Castillo, 2007; Vargas Sierra, 2008). Estos pasos no son necesariamente secuencias estrictas.

El desarrollo de la base de datos de saber terminológico que se propone (y que se explicará en el apartado siguiente con detalle) se vincula directamente con el paso (iii) *Diseño, construcción y explotación de corpus*, ya que esta constituirá un corpus de documentación terciaria para la construcción futura de un tesoro.

Además, hay que señalar que los tesauros representan la estructura nocional del dominio mediante un sistema de conceptos, entendido este último como un conjunto estructurado en el que se evidencian relaciones entre conceptos y, en el cual, cada uno está determinado por su posición respecto al resto (International Organization for Standardization, 2018).

Los tesauros y los métodos seguidos para su organización están relacionados con los procedimientos terminológicos para la construcción de sistemas conceptuales (Arntz & Picht, 1995); es decir, lo que desde la bibliotecología suele llamarse clasificación facetada (San Segundo Manuel, 1992), para la lingüística es una versión exclusivamente semántica del facetado, que se denomina sistema conceptual (5). En general, un sistema conceptual evidencia la percepción de los especialistas sobre su propia disciplina, los alcances de esta y las relaciones posibles para el ordenamiento de la terminología del área de especialidad –sean relaciones jerárquicas (lógicas o partitivas), secuenciales (cronológicas, causales, genéticas, entre otras) o pragmáticas (de contigüidad, temáticas, asociativas)(Arntz & Picht, 1995).

Desde una perspectiva clásica, las representaciones gráficas de los sistemas conceptuales son variadas: diagramas arborescentes, ontologías o, inclusive, sistemas con notaciones (Arntz & Picht, 1995); sin embargo, actualmente, las posibilidades de representación de un sistema conceptual son muy diversas, así como las herramientas para hacerlo son muy variadas (listas de encabezamientos, redes semánticas, mapas conceptuales, modelados tridimensionales, nubes de palabras, bases de conocimiento, esquemas informáticos, entre otros), lo que, en términos reales, facilita la representación del conocimiento de un dominio del saber en la web, por ejemplo.

En conclusión, el enfoque de este proyecto supone, en esta etapa, la construcción de un sistema simplificado que ordene los términos documentados por clases temáticas, autoridades que lo usan y tipos de términos (preferido o alternativo, y subdominio al que pertenece), basado en un sistema de organización del conocimiento semiformal (*Simple Knowledge Organization System*, SKOS) (W3C Working Group, 2009), que se verá con detalle en el siguiente apartado.

5. La estructuración de la base de datos de saber terminológico

Todas las ideas discutidas hasta el momento han determinado la estructura conceptual de la base de datos de saber terminológico, a partir de lo dicho, se considera que el desarrollo deberá ser

una base de datos que funja como un instrumento normalizador, alojado en línea, que represente al dominio de las artes espaciales en un sistema conceptual, y que sea armónico con sistemas de catalogación actuales.

En este apartado se ilustran las decisiones que se han tomado para el desarrollo conceptual de esta base de datos, considerando aspectos para su posible evolución hacia un tesoro construido a partir de un modelo semántico.

5.1. La base documental o controles terminológicos de la base de datos

Para la creación de una base de datos de saber terminológico es necesario contar con datos terminológicos. Este es el principal desafío de un proyecto como el que se plantea; es decir, el primer paso es documentar términos relacionados con el dominio de las artes espaciales que se relacionen con tipos de objetos y sus atributos físicos, materiales relacionados a estos, técnicas, estilos, periodos, entre otros.

Considerando los recursos en línea que constituyen un capital documental estructurado, de libre acceso y abierto para su reutilización, el proyecto comenzará con datos provenientes de: el *Tesoro de Arte & Arquitectura* y los *Tesauros-Diccionarios del patrimonio cultural de España*; además, se procederá a la digitalización de obras de consulta referidas al dominio para dar forma a la base de datos.

5.2. ¿Por qué una base de datos relacional para el proyecto?

Hooland y Verbough (2014) plantean la existencia de cuatro tipos de bases de datos: tabulares, relacionales, de estructuras jerárquicas (*meta-mark up languages*) y RDF (expresadas como tripletas) (6). Según ellos, el modelo tabular permite la atomización de datos y se estructura en archivos tan comunes como los “.csv” o “.xlsx”, lo que implica fácil manejo de información y accesibilidad técnica para el público en general.

Lamentablemente, un modelo tabular no acepta relaciones uno a muchos que, para el contexto del sector cultural, resultan muy frecuentes. Por ejemplo, hay casos de bienes culturales, como la Piedra de sol, disco monolítico de origen mexicana, hecho de basalto de olivino y exhibido en el Museo Nacional de Antropología de México, que se designan con múltiples títulos. Estos se organizan en tipologías reconocidas, por ejemplo, en estándares internacionales como VRA Core 4 (Visual Resources Association, 2007), a saber: nombres descriptivos, registrados, populares, traducidos a otras lenguas, alternativos, entre otros.

En ese caso, por su propia naturaleza, una base de datos tabular complica la sanción de múltiples títulos asociados a un solo objeto, así como la presencia de múltiples autores, técnicas, materiales, modos de uso, entre otros. Para ver más claramente esto se presenta la siguiente figura que sanciona información de Mexicana (Secretaría de Cultura, 2018):

ID	Título	Tipo de objeto	Fecha
0014137/0022948	Piedra del Sol	Pieza Arqueológica	1250-1521

Figura 1. Ficha abreviada de la “piedra de sol” en un modelo tabular

Como se puede ver en la Figura 1, en el campo título, solo se puede sancionar un nombre asignado al objeto en sí, sin que la información resulte ambigua, razón por la que se descarta este modelo, mientras que en una base de datos relacional (ejemplo de la Figura 2) se amplían las posibilidades que se tienen con un modelo tabular, y se permite vincular dos títulos a un mismo objeto mediante un identificador; incluso se puede especificar el tipo de título que se refiere y otras informaciones más.

ID	Título	Tipo de objeto	Fecha
0014137/0022948	1	Pieza Arqueológica	1250-1521

ID	Título	Tipo de título
1	Piedra del Sol	Nombre registrado
1	Calendario Azteca	Nombre popular

Figura 2. Ficha abreviada de la “piedra de sol” en un modelo relacional

Este modelo de bases de datos, además de permitir relaciones uno a muchos, como en el caso del título, requiere utilizar un modelo binario en SQL y, por lo tanto, implica un grado mayor de experiencia en el terreno de la construcción de bases de datos por parte de quien las crea (Hooland & Verborgh, 2014); sin embargo, hay algunas alternativas de software que facilitan esta tarea (MySQL, PostgreSQL, SQLite, HSQL, entre otros) y que hacen que este proyecto sea viable técnica y económicamente.

Por último, hay que señalar que, para el proyecto, una base de datos relacional es una buena solución inmediata, pero, a largo plazo, se considera que un desarrollo *ad hoc* como este podría caer en la obsolescencia conceptual y tecnológica. Sin embargo, en este momento, un modelo relacional es una alternativa para sistematizar la información que se pretende recoger y, en etapas siguientes, evolucionar el proyecto a un modelo semántico.

5.3. La estructura de la base de datos de saber terminológico

El modelo de organización del conocimiento “SKOS ha sido diseñado para proporcionar un modo de migrar a la Web Semántica sistemas de organización del conocimiento ya existentes con un coste bajo” (W3C Working Group, 2009), por lo que, desde este momento, se han incorporado elementos semánticos del modelo para, a mediano y largo plazo, migrar fácilmente la base de datos a un tesoro basado en un modelo semántico. El plan propuesto para ello se ve de forma esquemática en la siguiente figura:



Figura 3. Evolución de los datos asistemáticos a un modelo semántico

Como se puede ver en la Figura 3, el origen y la diversidad de los datos que se tienen, al seguir metodologías de elaboración distintas, son heterogéneos y requieren homogeneizarse, esto mediante la selección de datos comunes y la estructuración de ellos en un sistema de información. Lo anterior implica la transición de la primera etapa de documentación (globos de la izquierda) hacia la integración de la documentación diversa dentro de la base de datos (globos del centro). Al final, (globos de la derecha) se proyecta que el proceso terminará en la evolución del sistema hacia un modelo semántico.

5.4. Elementos mínimos de información de la base de datos de saber terminológico

El modelo de representación conceptual SKOS (W3C Working Group, 2009) es una alternativa simplificada de la norma ISO 25964 (International Organization for Standardization, 2018) para la construcción de tesauros y su interoperabilidad con otros instrumentos de control terminológico en línea. A partir de la revisión de estas dos iniciativas, la experiencia derivada del trabajo con catalogadores y una revisión de los estándares para la documentación de bienes culturales: VRA CORE 4 (Visual Resources Association, 2007) y CDWA (Harping & Baca, 2016), me permito concluir que los elementos mínimos con los que debe contar nuestro sistema de documentación terminológica son: (i) el término, (ii) la definición, (iii) un identificador único para cada noción y (iv)

la referencia a la autoridad que avala el uso de esta. Considerando estos cuatro elementos mínimos se determinó la estructura básica de la base de datos y de sus relaciones.

Según Cabré (1998), los términos designan los conceptos propios de cada disciplina; son unidades de forma y contenido que, como las unidades palabra, tienen una imagen acústica y una imagen conceptual, llamadas denominación y noción, respectivamente. Se considera que la denominación es la forma gráfica que toma cada término según las reglas ortográficas del español, por ejemplo; y que la noción está representada por las explicaciones, en forma de frases u oraciones, que se pueden encontrar en obras de consulta.

Como es sabido, una misma forma ortográfica se puede relacionar con significados distintos; a esta relación léxico-semántica se le conoce tradicionalmente como homonimia. Dicho de otra manera: “los homónimos se definen [...] como palabras distintas que poseen la misma forma” (Lyons, 1983, p. 49). Sin embargo, esta posibilidad de tener una multiplicidad de significados asociados a una misma forma también puede interpretarse como polisemia, entendida como “el fenómeno que consiste en que una palabra tenga, cuando se le considera en aislamiento [...] más de un significado” (Lara Ramos, 2006, p. 106).

En oposición al concepto de polisemia explicado antes, desde el punto de vista de la lexicografía, en terminología (Cabré i Castellví, 1998, p. 108),

[... the] semantic value of a term is established solely on the basis of its relationship to a specific conceptual system. Identifying a term as belonging to a special subject field involves placing it in a specific conceptual system, and as a result what in lexicography is considered polysemy, in terminology becomes homonymy.

En otras palabras, se podría decir que no hay términos polisémicos, sino que, en terminología, siempre se hablará de homonimia.

Teniendo claro el tratamiento de la polisemia, desde la terminología, y asumiendo que en la base de datos terminológica podría haber múltiples significados y que se deberán distinguir de su relación con la denominación (homonimia), el sistema asignará un identificador único a cada definición para funcionar como un elemento para desambiguar los conceptos e identificarse en la base de datos; con ello se establecerá que la denominación puede adquirir la misma forma ortográfica, pero que el concepto es único en el sistema de control terminológico. Naturalmente, se deberá incluir la definición y acompañarla con un elemento de información que refiera al tesoro en el que se registra; en este caso, se ha optado por sancionar las siglas del tesoro. Por último, se

relaciona la tabla principal (“Denominación”), en la que se harán las búsquedas de las denominaciones, y la tabla (“Nociones”) que concentra la información de la definición, las siglas del tesoro del que proviene y su identificador único. Estas dos tablas se relacionan mediante un identificador alfanumérico que adquiere la forma ortográfica del término. La lógica que subyace a esto se ilustra a continuación:

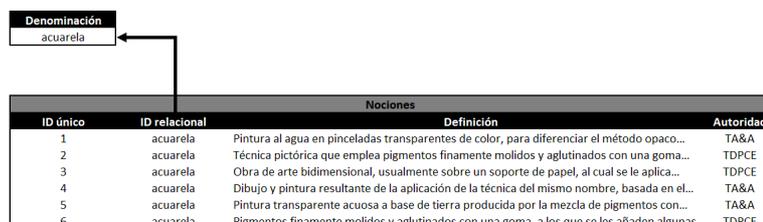


Figura 4. Elementos mínimos de información de la base de datos terminológica

Tal y como se ve en la Figura 4, hay dos tablas relacionadas; la primera funciona como un índice de búsqueda de denominaciones. Esta relación es posible mediante el elemento de información “ID relacional”. De esta forma, cuando el usuario busque el término *acuarela*, la búsqueda traerá como resultado seis nociones diferentes: dos asociadas a técnicas, dos a objetos y dos a materiales, las que se documentan, a su vez, en el TA&A y el TDPCE.

5.5. Otros elementos de información de la base de datos de saber terminológico

La iniciativa SKOS (W3C Working Group, 2009) señala que, en el plano de la representación ortográfica o transliteración del término, es conveniente hacer diferenciación entre la denominación preferida (skos:preflabel) y las denominaciones alternativas (skos:altlabel). Considerando lo anterior, a partir del ID de cada definición se asociarán denominaciones alternativas. Estos elementos de información amplían el modelo de esta forma:

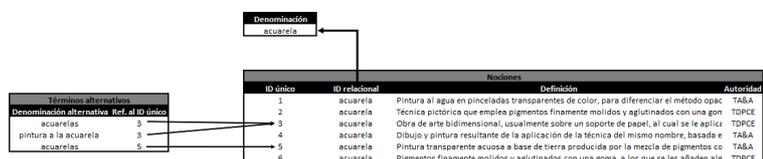


Figura 5. Relación de términos alternativos

Como se ve en la Figura 5, la denominación con ID único número 3, *acuarela*, tiene dos términos alternativos: *acuarelas* y *pintura a la acuarela*, ambos documentados en el TDPCE. Asimismo, la denominación documentada en el TA&A co-

respondiente al ID 5 solo tiene un término alternativo: *acuarelas*. De esta forma, el sistema dará cuenta de estas relaciones.

Finalmente, un vocabulario SKOS puede ser ampliado para adaptarse a las necesidades prácticas de comunidades de usuarios concretas, o combinadas con otros vocabularios de modelado (W3C Working Group, 2009); por lo tanto, en esta etapa de trabajo se ha considerado agregar al modelo tres informaciones que se consideran relevantes: (a) el tipo de noción al que hace referencia el concepto, es decir, si es una técnica, un tipo de objeto, un material, entre otros; (b) el subdominio de las artes espaciales en el que se ubica el concepto, esto es: pintura, escultura, fotografía, entre otras, y (c), si la denominación es usada en alguna institución custodia del patrimonio en México.

Considerando lo anterior, a la tabla "Nociones" se agregaron dos elementos de información más: el tipo de noción y el subdominio de este; ambos se incluyen directamente en la tabla, ya que cada uno de ellos se relaciona unívocamente con cada una de las nociones, y no se ordenan bajo relaciones uno a muchos, como es el caso de las denominaciones y las denominaciones alternativas. Esto se ve mejor a continuación (se resalta en color rojo):

Términos alternativos		Nociones					
Denominación alternativa	Ref. al ID único	ID único	ID relacional	Definición	Autoridad	Tipo de noción	Subdominio
acuarelas	3	2	acuarela	Pintura al agua en pinceladas transparentes de color, para diferenciar el método opaco.	TABA	Técnica	Pintura
pintura a la acuarela	3	3	acuarela	Técnica pictórica que emplea pigmentos finamente molidos y aglutinados con una goma.	TDICE	Técnica	Pintura
acuarelas	4	4	acuarela	Obra de arte bidimensional, usualmente sobre un soporte de papel, al cual se le aplica	TDICE	Objeto	Pintura
acuarelas	5	5	acuarela	Dibujo y pintura resultante de la aplicación de la técnica del mismo nombre, basada e	TABA	Objeto	Pintura
		6	acuarela	Pintura transparente acuosa a base de tierra producida por la mezcla de pigmentos co	TABA	Material	Pintura
			acuarela	Pigmentos finamente molidos y aglutinados con una goma, a los que se les añaden alq	TDICE	Material	Pintura

Figura 6. Suma al sistema del tipo de noción y subdominio

Control de autoridad (Institución)		Nociones					
Institución	Ref. den.	ID único	ID relacional	Definición	Autoridad	Tipo de noción	Subdominio
MNSC	acuarela	2	acuarela	Pintura al agua en pinceladas transparentes de color, para diferenciar el método opaco.	TABA	Técnica	Pintura
MNSC	acuarela	3	acuarela	Técnica pictórica que emplea pigmentos finamente molidos y aglutinados con una goma.	TDICE	Técnica	Pintura
MNSC	acuarela	4	acuarela	Obra de arte bidimensional, usualmente sobre un soporte de papel, al cual se le aplica	TDICE	Objeto	Pintura
MNSC	acuarela	5	acuarela	Dibujo y pintura resultante de la aplicación de la técnica del mismo nombre, basada e	TABA	Objeto	Pintura
MNSC	acuarela	6	acuarela	Pintura transparente acuosa a base de tierra producida por la mezcla de pigmentos co	TABA	Material	Pintura
MNSC	acuarela		acuarela	Pigmentos finamente molidos y aglutinados con una goma, a los que se les añaden alq	TDICE	Material	Pintura

Figura 7. Suma al sistema del control de autoridad (institución)

Por último, si la denominación es usada en alguna institución custodia del patrimonio en México, el desarrollo conceptual de la base de datos de saber terminológico dará cuenta de esto mediante una tabla adicional, "Control de autoridad (institución)"; esto con miras a que el sistema funja también como un control de autoridades (de institución). Los nombres de las instituciones se refieren por sus siglas y se relacionan con la tabla principal mediante una referencia directa de la denominación al elemento relacional: "Ref. den.". Esto se ve en la Figura 7 (se resalta en color

amarillo), queda garantía de uso de la denominación en instituciones de la memoria de México; esta información se recuperará de cuatro bases de datos del patrimonio existentes, a saber: Sistema Único de Registro Público de Monumentos y Zonas Arqueológicas e Históricas, Mexicana, Mediateca del INAH y Museos de México (las referencias completas se sancionan en la bibliografía). Lo que se puede ver en la ilustración es que en el Instituto Nacional de Antropología e Historia y en el Museo Nacional de San Carlos se usan regularmente estas denominaciones.

6. Resumen, conclusiones y futuras líneas de investigación

La demanda actual de datos estructurados, de información sistematizada y de colecciones de museos disponibles en entornos digitales determinan dinámicas específicas tocantes a la catalogación del patrimonio cultural y el control de sus vocabularios; por ello, la propuesta, aquí formalizada, de crear una base de datos de saber terminológico para el dominio parece tener un sustento en necesidades concretas de una comunidad de práctica, a saber: especialistas de museos, archivos, bibliotecas, galerías, así como historiadores del arte.

Aunque se han identificado algunos proyectos de control terminológico destacados, se presentaron razones por las cuales el planteamiento de un proyecto de tesoro para México tendría sentido, es decir: resolvería la ambigüedad en las definiciones, la variación del tecnoléxico desde el punto de vista geográfico, la necesidad de localización de las definiciones, la armonización de trabajos existentes con sistemas y estándares actuales de catalogación del patrimonio, entre otros. En consecuencia, se plantea el desarrollo de una base de datos que documente el uso de los términos dentro de la comunidad de práctica de México, y que nos permita reconocer diferencias de uso geográficas, temporales, de grado de especialización y usos jergales que se traducirán en un mejor tesoro para México.

En cuanto al desarrollo conceptual de la base de datos de saber terminológico, paso previo a la construcción del tesoro, se han discutido una serie de aspectos teóricos que determinaron la estructura de la base de datos, todos ellos provenientes de disciplinas humanas como la biblioteología, la terminología, la historia del arte, entre otras. Este proceso se puede resumir en cuatro puntos específicos que se recapitulan a continuación, ello para que el proceso pueda replicarse en otros proyectos de alcance semejante:

a. *Selección de la documentación.* Esta selección debe obedecer a criterios de disponibilidad documental, posibilidades de uso abierto de los datos

o bajo permisos, representatividad de la variante de la lengua que se pretende describir, y, por último, a criterios de selección de las obras de consulta por parte de especialistas en el dominio.

b. *Determinar el tipo de base de datos que se va a desarrollar.* Esto responde a las capacidades tecnológicas, al grupo de trabajo y recursos materiales y humanos con los que cuente el proyecto.

c. *Establecer los elementos mínimos requeridos.* Este proceso es el resultado de la revisión del estándar SKOS CORE y la norma ISO 25964. Con ello se propondrá un modelo armónico con las normas y estándares internacionales.

d. *Pensar en las necesidades de la base de datos y ampliar los elementos de información que sea pertinente incluir.* En este paso, se pueden ampliar los elementos de información al modelo que se ha propuesto en el paso anterior. Esto se logra con un diagnóstico de las necesidades de los usuarios potenciales de la base de datos de saber terminológico. Este análisis es responsabilidad del grupo de trabajo.

En cuanto al futuro de la base de datos y su sistema conceptual, en este momento el diseño solo permitiría dar cuenta de las diferencias entre términos preferidos (skos:preflabel) y los términos alternativos (skos:altlabel). Sin embargo, se considera que en un futuro será importante incluir relaciones jerárquicas de términos genéricos (skos:narrower) y términos específicos (skos:broader), así como relaciones entre términos asociados (skos:related) y, por último, conviene incluir ejemplos de uso que remitan a documentación secundaria (skos:example) (W3C Working Group, 2009).

Hay que decir que lo anterior no tiene sentido en esta etapa de trabajo, ya que consideramos incrementar los registros de la base de datos hasta el punto en que logremos documentar la gran mayoría de los términos de las bases de datos del patrimonio que ya se han referido antes (Sistema Único de Registro Público de Monumentos y Zonas Arqueológicas e Históricas, Mexicana, Mediateca del INAH y Museos de México); lo anterior en virtud de que solo así podríamos dar un panorama más completo de las relaciones léxico-semánticas de todo el dominio del saber.

En cuanto a futuras líneas de trabajo, esta base de datos constituirá la base empírica que permitirá la consolidación de investigaciones sobre el *clustering* de información, la generación automática de taxonomías, la determinación de una metalengua definitoria para las artes espaciales, entre otras.

Por último, se reconoce que el diseño de un modelo de datos como el que se presenta podría ser considerado como un trabajo poco innovador; sin

embargo, se advierte que, en el contexto del sector cultural mexicano, no hay antecedentes concretos sobre iniciativas como esta. Por lo tanto, la innovación en sí es el inicio de una línea de trabajo en un sector en el que no se registra antecedente alguno. También, se hace una última aclaración: este trabajo es una etapa intermedia necesaria para consolidar una base empírica que servirá para investigaciones futuras.

Notas

- (1) Desde el punto de vista de la teoría del arte, las artes podrían clasificarse en artes espaciales, temporales y espacio-temporales. El proyecto que aquí se presenta se centra en artes espaciales como la pintura, escultura, grabado, fotografía, entre otras. Para entender mejor esta clasificación de las artes basada en los conceptos tiempo y espacio, se recomienda el libro de Escobar (2015).
- (2) En la publicación de Recio Crespo (2016), desde la experiencia puntual de distintos colaboradores del Proyecto de Tesoros del Patrimonio Cultural de España, se enfatiza la necesidad de integración de tesoros en proyectos de más grande envergadura, la implementación de normas y estándares para la elaboración de tesoros, y la importancia de los tesoros para la cosecha de información sobre el patrimonio cultural, que, en mi opinión, son iniciativas dignas de ser usadas como modelos referenciales.
- (3) Un *mojigango* es una figura antropomórfica grotesca, de cartón pintado, que muchas veces representa políticos y otros personajes de la vida pública. Esta realidad tiene origen en el estado de Colima, en México, y ha trascendido como objeto patrimonial en museos regionales y nacionales de culturas populares en nuestro país.
- (4) En este artículo no se explicará ni ahondará en una propuesta metodológica sobre cómo se pretende construir un facetado que refleje cabalmente el conocimiento representado en el proyecto, esto en virtud de que el objetivo del presente no es ello. Para ello, se recomienda la revisión de Molina Salinas & Franco Trujillo (2018).
- (5) Desde luego que existen algunas diferencias, sobre todo desde el punto de vista de la organización de la información, mismas que dan pie a la generación de clases o materias, pero, para efectos de esta propuesta, no es relevante ahondar en ellas, en este momento, y se refiere al libro de San Segundo Manuel (1992).
- (6) Sobre los modelos jerárquicos y de RDF, Hooland y Verbough (2014) explican claramente, en el capítulo 2: "Modeling", las ventajas de cada uno. Debido a que lo que se plantea con este trabajo es formalizar una base de datos relacional que, en un futuro, sea escalable a este tipo de modelos complejos, se refiere a él para estos pormenores.

Referencias

- Alpizar Castillo, Rodolfo (2007). *¿Cómo hacer un diccionario científico-técnico?*. Buenos Aires: Ed. Memphis, 2007.
- Arano, Silvia (2005). Los tesoros y las ontologías en la Biblioteconomía y la Documentación. // *Hipertext.Net. Anuario Académico Sobre Documentación Digital y Comunicación Interactiva*, 3 (2005). <https://www.upf.edu/hipertextnet/numero-3/tesoros.html> (2019-05-01).
- Arntz, Reiner; Picht, Heribert (1995). *Introducción a la terminología*. Madrid: Fundación Germán Sánchez Ruipérez, 1995.
- Barité Roqueta, Mario Guido (2014). El control de vocabulario en la era digital: revisión conceptual. // *Scire*. 20(1)

- (2014). <http://ibersid.eu/ojs/index.php/scire/article/view/4196/3766> (2019-05-01).
- Cabrè i Castellví, María Teresa (1998). *Terminology. Theory, Methods, and Application*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1998.
- Centro de Documentación de Bienes Patrimoniales. (n.d.). *Tesaurus de Arte & Arquitectura*. <http://www.aatespanol.cl/> (2019-05-01).
- Combi, Mariella (2016). Culture and Technology: An Analysis of Some of the changes in Progress - Digital, Global and Local Culture. // Borowieki, Karol Jan; Forbes, Neil; Fresa, Antonella (eds.). *Cultural Heritage in a Changing World*. Gewerbestrasse: Springer Open, 2016. 3-16.
- Coseriu, Eugen (1985). *El hombre y su lenguaje. Estudios de teoría y metodología lingüística*. Madrid: Ed. Gredos, 1985.
- Doerner, Max (2005). *Los materiales de pintura y su empleo en el arte*. Barcelona: Editorial Reverté, S.A., 2005.
- Elings, Mary; Waibel, Günter (2007). Descriptive Standards and Metadata Sharing across Libraries, Archives and Museums. // *First Monday*. 12:3. <https://doi.org/10.5210/firstmonday.12i3.1628> (2019-05-01).
- Escobar Rodríguez, Irma (2005). *Conocimientos fundamentales para la formación artística*. México: UNAM.
- Fischer, Andreas (2004). The Notional Structure of Thesauruses. // Kay, Christian; Smith, Jeremy (eds.). *Categorization in the History of English*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2004. 41-58.
- Harping, Patricia (2010a). Development of the Getty Vocabularies: AAT, TGN, ULAN, and CONA. // *Art Documentation: Journal of the Art Libraries Society of North America*. 29:1 (2010a) 67-72. <https://doi.org/10.1086/adx.29.1.27949541> (2019-05-01).
- Harping, Patricia (2010b). *Introduction to Controlled Vocabularies. Terminology for Art, Architecture, and Other Cultural Works*. Los Angeles: Getty Research Institute, 2010.
- Harping, Patricia; Baca, Murtha (2016). Categories for the Description of Works of Art (CDWA). // Getty Research Institute. (2016) http://www.getty.edu/research/publications/electronic_publications/cdwa/index.html (2019-05-01).
- Hooland, Seth van; Verborgh, Ruben (2014). *Linked data for libraries, archives and museums: how to clean, link and publish your metadata*. London: Facet Publishing, 2014.
- INEGI. (2015). *Encuesta Intercensal 2015 Estados Unidos Mexicanos*. México, D.F.: Instituto Nacional de Estadística y Geografía, 2015.
- International Organization for Standardization. (2018). ISO/DIS 1087:2018 (E). *Terminology work - Vocabulary*. Genève: International Organization for Standardization, 2018.
- Instituto Nacional de Antropología e Historia. (2018). *Mediateca del INAH // Instituto Nacional de Antropología e Historia | Repositorio Institucional*. https://mediateca.inah.gob.mx/islandora_74/ (2019-05-01).
- Instituto Nacional de Antropología e Historia (2018). *Sistema Único de Registro Público de Monumentos y Zonas Arqueológicas e Históricas*. // *Página del Registro Público de Monumentos y Zonas Arqueológicas e Históricas*. <http://www.registropublico.inah.gob.mx/index.php/autenticacion/autenticacion> (2019-05-01).
- Lara Ramos, Luis Fernando (2006). *Curso de lexicología*. México, D.F.: El Colegio de México, 2006.
- Lyons, John (1983). *Lenguaje, significado y contexto*. Barcelona/Buenos Aires: Ediciones Paidós, 1983.
- Ministerio de Cultura y Deporte. (n. d.). *Tesoros - Diccionarios del patrimonio cultural de España - Acceso a los tesauros*. // *Web del Ministerio de Cultura y Deporte*. <http://tesoros.mecd.es/tesauros/tesauros> (2019-05-01).
- Molina Salinas, Claudio (2017). Desarrollo de un diccionario numismático basado en estándares internacionales para la catalogación de objetos culturales. // *Revista Numismática HÉCATE*. 4 (2017). <http://hdl.handle.net/10201/55652> (2001-01-22).
- Molina Salinas, Claudio; Franco Trujillo, Erik Daniel (2018). Una metodología para la construcción de un sistema conceptual para un tesoro de las artes espaciales en México. // *Termin@lia*. 18. <https://doi.org/10.2436/20.2503.01.124> (2001-01-22).
- Muñiz Castro, Emilio (2004). La normalización terminológica del español y los diccionarios especializados. González, Luis; Hernández, Pollux (eds.). *Las palabras del traductor. Actas del II Congreso «El español, lengua de traducción»*. Toledo, 20 y 21 de mayo. Bruselas: ESLETRA, 2004. 221-242.
- Naumis Peña, Catalina (2007). *Los tesauros documentales y su aplicación en la información impresa, digital y multimedia*. México, D.F.: Universidad Nacional Autónoma de México/Alfagrama Ediciones, 2007.
- Recio Crespo, Miguel Ángel (2016). *El lenguaje sobre el patrimonio. Estándares documentales para la descripción y gestión de colecciones*. Madrid: Ministerio de Educación, Cultura y Deporte.
- Rodríguez Ortega, Nuria (2016). Construcción y uso de terminologías, categorías de descripción y estructuras semánticas vinculadas al patrimonio en la sociedad global de datos. // Recio Crespo, Miguel Ángel (ed.). *El lenguaje sobre el patrimonio. Estándares documentales para la descripción y gestión de colecciones*. Madrid: Ministerio de Educación, Cultura y Deporte, 2016. 115-130.
- San Segundo Manuel, Rosa. (1992). *Teoría e historia de la clasificación bibliotecaria en España, siglos XIX y XX*. Universidad Complutense de Madrid. Tesis doctoral, 1992.
- Sagredo Fernández, Félix; Espinosa Temiño, María Blanca (2000). *Del libro, al libro electrónico-digital*. // *Cuadernos de Documentación Multimedia*. 9 (2000).
- Secretaría de Cultura (2018). *Mexicana. Repositorio del Patrimonio Cultural de México // Página del Repositorio del Patrimonio Cultural de México*. <https://mexicana.cultura.gob.mx/es/repositorio/home> (2001-01-22).
- Secretaría de Cultura. (2018). *Museos de México. // Página de la secretaría de Cultura*. <http://museosmexico.cultura.gob.mx/> (2001-01-22).
- The Getty Research Institute (n.d.). *The Art & Architecture Thesaurus On Line // The Getty Research Institute Web Page*. <http://www.getty.edu/research/tools/vocabularies/aat/index.html> (2019-05-01).
- Vargas Sierra, Chelo (2008). La sistematización terminográfica: una propuesta metodológica para la elaboración de diccionarios traductológicos. // *Actas del X Simposio Iberoamericano de Terminología: Montevideo, 6 al 9 de noviembre*. Montevideo: RITerm, Red Iberoamericana de Terminología. 1-25. <http://rua.ua.es/dspace/handle/10045/13212>.
- Visual Resources Association. (2007). *VRA Core 4.0 Element Description*. // *VRA CORE Official Web Site*. https://www.loc.gov/standards/vracore/VRA_Core4_Element_Description.pdf (2001-01-22).
- W3C Working Group. (2009). *Manual de SKOS (Simple Knowledge Organization System, Sistema Simple de Organización del Conocimiento)*. // *W3C Working Group Note*. <http://skos.um.es/TR/skos-primer> (2001-01-22).

Enviado: 2019-05-04. Segunda versión: 2019-06-18.

Aceptado: 2019-07-30.
