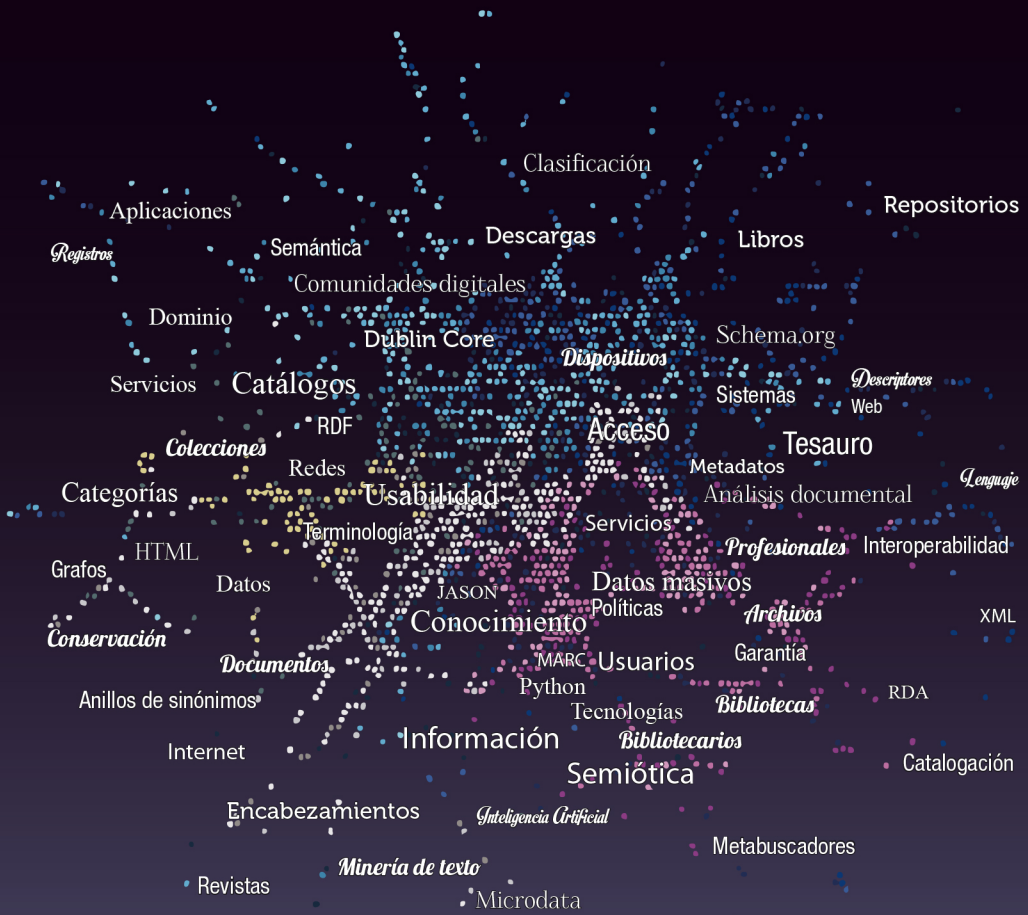


Organización del conocimiento

Bibliotecología, Terminología e Información



Catalina Naumis Peña
Adriana Suárez Sánchez
Coordinadoras



Z1006
O743

Organización del conocimiento : bibliotecología, terminología e información /
coordinadoras Catalina Naumis Peña, Adriana Suárez Sánchez. –Primera edición.
– Ciudad de México : Universidad Nacional Autónoma de México, Instituto de
Investigaciones Bibliotecológicas y de la Información, 2025.
xi, 240 páginas. – (Sistematización de la información documental)

Este libro fue producido en el marco del Proyecto PAPIIT IT 400122 Sistema
de información terminológica sobre bibliotecología, archivología e información.

ISBN: 978-607-587-131-8

1. Organización del conocimiento. 2. Terminología bibliotecológica.
I. Naumis Peña, Catalina, editor. II. Suárez Sánchez, Adriana, editor. III. serie.

Este libro fue producido en el marco del Proyecto PAPIIT IT 400122 *Sistema de
información terminológica sobre bibliotecología, archivología e información.*

Diseño de portada: Carlos Papaqui Landeros

Primera edición: febrero de 2025

D. R. © UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
Instituto de Investigaciones Bibliotecológicas y de la Información
Circuito Interior s/n, Torre II de Humanidades,
pisos 11, 12 y 13, Ciudad Universitaria, C. P. 04510,
Alcaldía Coyoacán, Ciudad de México

ISBN: 978-607-587-131-8

Esta edición y sus características son propiedad de la Universidad Nacional Autónoma
de México. Prohibida la reproducción total o parcial por cualquier medio sin la
autorización escrita del titular de los derechos patrimoniales.

Publicación dictaminada
Hecho en México

Contenido

Presentación	vii
<i>Birger Hjørland</i>	

Prefacio	ix
<i>Catalina Naumis Peña</i>	

I.

PERSPECTIVAS DE EXPERTOS

La transformación conceptual y terminológica sobre catalogación en el siglo XXI	3
<i>Catalina Naumis Peña</i>	

Distribución de relaciones léxicas como indicador de propiedades estructurales en vocabularios controlados: el caso de la sinonimia	21
<i>Diego Ferreyra</i>	

La enseñanza de la construcción de tesauros: experiencias y reflexiones	39
<i>Francisco Javier García Marco</i>	

II.

APLICACIONES PRÁCTICAS

La representación temática facetada: análisis y propuesta de conversión y representación de encabezamientos de materia a FAST	63
<i>Noé Ángeles Escobar</i>	

Los sistemas de anotación como instrumento de organización del conocimiento	87
<i>Jesús Tramullas</i>	

Hacia una sistematización de la terminología de las ediciones: una contribución desde la organización del conocimiento a la industria editorial. 103

Mario Barité

Carolina Saravia Rebollo

Consolidando un ecosistema de documentación del patrimonio histórico y cultural de la UNAM: propuesta metodológica *ad hoc* para el desarrollo de un microtesauro. 121

Claudio Molina Salinas

III.

PERSPECTIVAS INNOVADORAS

Gestión de datos digitales: desafíos de la organización de información. 141

Lourdes Feria Basurto

Elizabeth Nashielli Márquez Sánchez

Repositorios institucionales, redes y organización del conocimiento: un estado de la cuestión en Hispanoamérica. 163

Gema Bueno-de-la-Fuente

María del Carmen Agustín-Lacruz

IV.

ORGANIZACIÓN DEL CONOCIMIENTO DIGITAL

Mapas tópicos y ontologías terminológicas: características y tendencias en la organización temática de colecciones digitales. 195

Adriana Suárez Sánchez

Las relaciones terminológicas entre los sistemas de información y los sistemas informáticos. 211

Hugo Alberto Guadarrama Sánchez

La organización de la información en la era digital. 227

Blanca Estela Sánchez Luna

Consolidando un ecosistema de documentación
del patrimonio histórico y cultural de la UNAM:
propuesta metodológica *ad hoc* para el desarrollo
de un microtesauro

*Consolidating an ecosystem of documentation of the historical
and cultural heritage of the UNAM: ad hoc methodological
proposal for the development of a microthesaurus*

Claudio Molina Salinas
Instituto de Investigaciones Estéticas
Universidad Nacional Autónoma de México

INTRODUCCIÓN

En los últimos diez años, desde una perspectiva GLAM (*Galleries, Libraries, Archives and Museums*),¹ académicos y personal de algunos museos, facultades e institutos de la Universidad Nacional Autónoma de México (UNAM) emprendimos un trabajo enfocado en consolidar la documentación del patrimonio histórico y cultural que custodiamos en la Universidad, hacer trabajo de investigación en torno a este, y tomar acciones como utilizar un sistema de gestión de colecciones de licencia abierta (CollectiveAccess), que permite la gestión de estas y su publicación en línea;² implementar buenas prácticas de catalogación, adoptando procedimientos y estándares internacionales como VRA Core 4;³ implementar reglas de catalogación de bienes de interés cultural

1 Marina Salse *et al.*, “GLAM Metadata in Museums and University Collections: A State-of-the-art (Spain and other European Countries)”, *Global Knowledge, Memory and Communication* 73, 4/5, 2024, 478.

2 “CollectiveAccess”, *collectiveaccess.org*, 2022, <https://www.collectiveaccess.org/>.

3 Visual Resources Association, “VRA Core 4.0 Element Description”, Visual Resources, 2007, https://www.loc.gov/standards/vracore/vra_Core4_Element_Description.pdf

(BIC)⁴ como Cataloguing Cultural Objects (CCO),⁵ entre otros. Estas son las dependencias universitarias representadas en esta comunidad:

- Antigua Academia de San Carlos (ASC), Facultad de Artes y Diseño (FAD).
- Archivo Fotográfico Manuel Toussaint (AFMT), Instituto de Investigaciones Estéticas (IIE).
- Centro de Documentación Arkheia (CDA).
- Centro de Interpretación Xaltilolli (CIX), Centro Cultural Universitario Tlatelolco (CCUT).
- Dirección General de Repositorios, UNAM (DGRU-UNAM).
- Museo Universitario de Arte Contemporáneo (MUAC).
- Museo Universitario del Chopo (MUCH).
- y la Unidad de Información para las Artes (UNIARTE), IIE.

Luego de varios años de arduo trabajo y de cosechar notables avances individuales en la documentación de nuestras colecciones, debido a que encontramos una amplia dispersión en el uso de los términos descriptores de estas, a principios del año 2023 nos propusimos desarrollar un microtesauro de uso común que sea, a la vez, un control terminológico *ad hoc*, un sistema local de organización del conocimiento y un instrumento de desambiguación conceptual para los catalogadores de nuestras dependencias.

En este capítulo se presenta el proceso de creación y publicación de un microtesauro que recoge términos relativos al tipo de objeto, únicamente; también se ahonda en algunas decisiones tomadas al momento de: primero, integrar la información de todas nuestras bases de datos (BD); segundo, normalizar las terminologías que se usan en los distintos museos y centros de investigación, poniendo particular atención a los que denominan tipos de objetos; tercero, definir los términos preferidos y los alternativos; cuarto, explicar qué modelo de definición se ha adoptado y la estrategia de redacción

4 En España se usa el término legalmente para denominar a los bienes muebles o inmuebles del patrimonio cultural de la nación, y con la publicación en línea de “Mexicana: el repositorio cultural de México”, se ha incorporado a nuestro contexto este concepto para describir al conjunto de bienes de la nación, pese a no tener implicación legal alguna. Ministerio de Educación, Cultura y Deporte, “Definición de Bien de Interés Cultural”, Página del Ministerio de Educación, Cultura y Deporte, *s/f*, <https://www.cultura.gob.es/cultura/patrimonio/bienes-culturales-prottegidos/definicion.html>; José Manuel Morales-del-Castillo, Pedro Ángeles Jiménez, y Claudio Molina Salinas, “Mexico’s Tradition and Culture Entering the Digital Age: The Mexican Cultural Heritage Repository Project”, *Heritage 2*, 2019, 356-65.

5 Murtha Baca *et al.*, eds., *Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images* (Chicago: American Library Association, 2006).

de estas; quinto, revisar las definiciones y delimitar las relaciones jerárquicas entre los términos; y, por último, se argumentan las razones por las que hemos optado por utilizar un *software* y solución tecnológica como TemaTres para la publicación del microtesauro.

1. UN ECOSISTEMA DE CATALOGACIÓN DE BIC EN LA UNAM

En el año 2012, Renato González Mello (director del IIE), en coordinación con funcionarios de dependencias de órganos desconcentrados del Gobierno Federal (Centro Nacional de Conservación y Registro del Patrimonio Artístico Mueble del Instituto Nacional de Bellas Artes y Literatura; Dirección General de Sitios y Monumentos del Patrimonio Cultural, y Registro Público de Monumentos y Zonas Arqueológicas e Históricas, ambos del Instituto Nacional de Antropología e Historia), junto con autoridades de dependencias universitarias (Instituto de Investigaciones Bibliotecológicas y de la Información; y Unidad de Informática para la Biodiversidad (UNIBIO) del Instituto de Biología (IB), actualmente la DGRU-UNAM), organizaron un encuentro intensivo para conocer y discutir sobre los métodos usados en México para la estandarización y organización de registros de BIC.⁶

En seguimiento a los trabajos de esta reunión, en el IIE se fundaría la UNIARTE, institucionalizada hasta el año 2016, que es, a su vez, un punto de encuentro, un seminario permanente y una unidad de información que tiene como finalidad integrar todos los datos organizados de nuestro Instituto, así como impulsar las mejores prácticas de catalogación y publicación de datos científicos sobre el arte y llevarlos, en su momento, hasta un formato de datos abiertos vinculados.^{7 y 8}

En la UNIARTE, además de plantearnos discusiones teóricas y revisión de diferentes estándares y metodologías, adoptamos un modelo de catalogación del patrimonio utilizando el *software* de licencia abierta CollectiveAccess (creado por Whirl-i-Gig), que sirve para gestionar y difundir colecciones de museos, archivos y bibliotecas.⁹ Usando este *software*, en 2015, en la UNIARTE acom-

6 Comunicación personal con Renato González Mello, exdirector del IIE-UNAM, abril de 2024.

7 Comunicación personal con Pedro Ángeles Jiménez y Gabriela Betsabé Miramontes Vidal, uniarte, abril de 2024.

8 Unidad de Información para las Artes (UNIARTE), “Unidad de Información para las Artes, Instituto de Investigaciones Estéticas”, Página web del Instituto de Investigaciones Estéticas, 2020, <http://www.esteticas.unam.mx/uniarte>.

9 Rubén Alcaraz Martínez, “CollectiveAccess, un sistema de gestión y difusión de colecciones de museos, archivos y bibliotecas”, *BiD: textos universitarios de biblioteconomía e documentación*, 33, 2014.

pañamos la migración de la antigua BD del AFMT a una instalación de CollectiveAccess; a finales del 2016 comenzaríamos a administrar tanto una BD del patrimonio de México con más de 20 000 registros y, debido a los sismos del 2017, se crearía otra BD de BIC afectados por desastres naturales, que contiene más de 2 000 registros.¹⁰

En los últimos ocho años otros miembros de nuestro seminario también adoptarían CollectiveAccess y publicarían exitosamente sus colecciones; por ejemplo: en 2017, la ASC publicaría el catálogo digital de sus colecciones;¹¹ desde 2021, las colecciones del MUAC se han ido documentando en una instalación de CollectiveAccess y están próximas a publicarse en 2024; y, en mayo de 2024, se integró a las actividades de nuestro seminario el MUCH, que publicó, en 2021, *Desobediente: Archivo digital del Museo Universitario del Chopo*.¹²

Por otra parte, el CIX no usa *CollectiveAccess*, sino Tainacan,¹³ como otra alternativa de *software* de código abierto en el que se pueden administrar, gestionar y publicar colecciones digitales;¹⁴ asimismo, en el CDA comenzarán próximamente con el proceso de publicación de sus colecciones (aún no definen el *software* gestor para las mismas).¹⁵ Estos dos casos sirven para aclarar que CollectiveAccess no es “la solución definitiva”, ya que reconocemos más importante para nuestra comunidad el mantener la implementación de buenas prácticas de documentación, que el uso de un *software*.

Claramente, aún hay mucho trabajo por hacer, no obstante, desde enero de 2023 nos hemos propuesto resolver una necesidad de forma colectiva, es decir: contar con un microtesauro para normalizar el uso de la terminología con la que describimos nuestras colecciones, en el que, en principio, incluiremos únicamente a tipos de objetos (caso práctico descrito en este capítulo) y, en etapas posteriores, continuaremos con materiales, técnicas, etcétera. A continuación se explicarán los pasos de la metodología desarrollada *ad hoc* para el microtesauro y se ilustrarán algunas decisiones tomadas para su construcción.

10 Comunicación personal con Pedro Ángeles Jiménez, coordinador de uniarte, abril de 2024.

11 <https://coleccionessancarlos.fad.unam.mx/>

12 <https://archivodesobediente.chopo.unam.mx/>

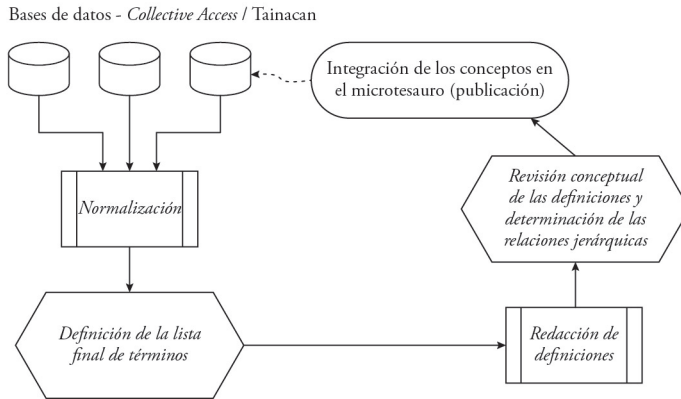
13 Tainacan es un “open source, powerful and flexible repository platform for WordPress. Manage and publish your digital collections as easily as you publish a post on your blog while having all professional repository platform tools available”. “Tainacan Wiki”, The Tainacan Project Web’s page, 2020, <https://tainacan.github.io/tainacan-wiki/#/?id=tainacan-wiki>

14 Comunicación personal con Lucía Sánchez de Bustamante, coordinadora del CIX, abril de 2024.

15 Comunicación personal con Pedro Ángeles Jiménez, coordinador de UNIARTE, abril de 2024.

2. FLUJO DE TRABAJO PARA LA CONSTRUCCIÓN DEL MICROTESAURO

El método seguido para la construcción del microtesauro consta de cinco pasos diferenciados, consecutivos, que llevan a una o varias unidades léxicas a través de ciertos procesos de cotejo y documentación hasta su publicación en línea. Los pasos propuestos son: (1) integración de las BD y normalización de los términos, (2) definición de la lista final de términos, (3) redacción de definiciones, (4) revisión conceptual de las definiciones y determinación de las relaciones jerárquicas, y (5) integración de los conceptos en el microtesauro (publicación en TemaTres). Lo anterior se puede ver gráficamente en el siguiente esquema:



1. Propuesta metodológica para la creación de un microtesauro (diagrama de flujo).


Adoptar este flujo de trabajo acorta notablemente los tiempos de revisión, dictaminación y publicación del microtesauro. A continuación, explicaré cada uno de estos momentos del diagrama anterior y señalaré algunas decisiones tomadas para un par de casos excepcionales.

2.1. Integración de las BD y normalización de los términos

Para el primer paso de la construcción del microtesauro hay que concentrar los términos usados en las BD o fuentes de información disponibles para el proyecto y también normalizarlos. De tal manera que integramos los términos o descriptores sancionados en el elemento de información “tipo de objeto” de todas las BD en un solo archivo csv, anotando debidamente la fuente de

procedencia de cada uno de los registros. Este proceso dio forma a un listado terminológico con 718 691 ocurrencias (*tokens*) y 991 tipos (*types*).¹⁶ Como se verá más adelante, hacer esta distinción entre ocurrencias y tipos, y registrar la fuente de procedencia de los datos fue relevante para la toma de decisiones y metodologías a lo largo de todo el flujo de trabajo.

En el listado de 718 691 ocurrencias nos encontramos casos ilegibles debido a problemas con la codificación, así como algunos errores de captura (con espacios al inicio o al final de la unidad léxica o formas con espacios consecutivos), términos con “errores de dedo” o escritos en su totalidad en mayúsculas. Para resolver esto, el terminólogo del seminario trató estos datos con OpenRefine,¹⁷ *freeware* usada para limpiar y transformar BD desordenadas. En la imagen siguiente, a la izquierda, vemos los datos antes de ser procesados en OpenRefine y, a la derecha, el resultado de las transformaciones hechas con la herramienta



657087	VISTA		
657088	VISTA		
657089	VISTA		
657090	VISTA		
657091	VISTAS		
657092	VISTAS		
657093	vitral		
657094	Zona arqueológica		
657095	Zona arqueológica		
657096	zona arqueológica		
657097	Zona arqueológica		
657098	Zona arqueológica		
657099	Zona arqueológica		
657100	Zona arqueológica		
657101	Zona arqueológica		
657102	Zona arqueológica		
657103	Zona arqueológica		
657104	Zona arqueológica		
657105	Zona arqueológica		

O352			
	A	B	C
334	universidad		
335	universidad estatal		
336	urna		
337	utensilio		
338	vasija		
339	vasija (contenedores por forma)		
340	vasija para beber		
341	vaso		
342	vasp		
343	via cotidiana		
344	vida cotidiana		
345	vida cotidiana/tipos populares		
346	vista		
347	vistas		
348	vitral		
349	zona arqueológica		

2. Listado de términos del AFMT antes y después del refinamiento de datos hecho con OpenRefine.

Como se puede observar en el resultado del refinamiento mostrado en la imagen anterior, aún persisten errores ortográficos o “de dedo” como “via cotidiana” o “vasp”. En este caso, se corrigieron manualmente. También se eliminaron de la lista voces como “universidad estatal”, “vida cotidiana” y “vida cotidiana/tipos populares”, ya que no son unidades morfológicamente fijas deberían estar en otro elemento de información, y son un error de captura por parte del documentalista, respectivamente. Asimismo, descartamos los cuali-

16 “Ocurrencia” es cada una de las unidades léxicas que aparecen en el listado de términos que concentramos y “tipo” es la unidad léxica registrada eliminando sus repeticiones. Luis Fernando Lara Ramos, Curso de lexicología (México, D.F.: El Colegio de México, 2006), 155-56.

17 “OpenRefine”, The Open Refine Web’s Page, 2012, <https://openrefine.org/>.

ficadores, como en el caso: “vasija (contenedores por forma)”, conservando únicamente “vasija”, ya que la gran mayoría de las terminologías de nuestras BD no utilizan este recurso de desambiguación y porque nosotros optamos por enumerar las terminologías para resolver la ambigüedad.

Finalmente, terminamos la normalización de los términos ajustándonos a los criterios usados para definir vocablos, según la usanza de la lexicografía española, llevando plurales a singulares y femeninos a masculinos cuando fue el caso.¹⁸ Además, revisamos los términos cotejando las reglas ortográficas del español y contra algunos diccionarios especializados.¹⁹ Todos estos procesos derivaron en una lista de 991 tipos de objetos con los que trabajamos en la etapa siguiente de esta metodología.

2.2. Definición de la lista final (o cercana) de términos

Para definir una lista de términos validados por especialistas es necesario reflexionar sobre la autenticidad de estos en un seminario y, simultáneamente, hacer un análisis cuantitativo que defina la frecuencia de los términos para establecer un orden de definición de estos, así como establecer los términos preferidos (descriptores) y los alternativos (no descriptores).

Para ello, previo a algunas de las reuniones del seminario permanente, el terminólogo del grupo convirtió la lista de 991 candidatos a términos en un archivo .DOCX, se puso en contacto con los colegas de los museos y dependencias universitarias involucradas en el seminario y les hizo llegar solo los términos que utilizan para catalogar sus colecciones. Vía correo electrónico y en el seminario se socializaron las dudas que resultaron de estas revisiones, se aclararon algunos casos que no fueron del todo claros y, por último, se generó una lista final de 715 términos aceptados de forma colegiada.

Por otro lado, podría pensarse que para un proyecto de microtesauro como este, precisar el orden de redacción de las definiciones de los términos es irrelevante, ya que hay que definirlos todos; sin embargo, debido a la naturaleza del trabajo del seminario, la forma inmediata en la que se publican las colecciones en CollectiveAccess y Tainacan, el modo de publicación del microtesauro en el *software* TemaTres y la necesidad de contar con el microtesauro, hemos optado

18 Luis Fernando Lara Ramos, *Teoría del diccionario monolingüe* (México, D.F.: El Colegio de México, 1997), 118-119, 170.

19 Algunos recursos consultados son los Tesauros-Diccionarios del patrimonio cultural de España, Art & Architecture Thesaurus, Tesoro de la UNESCO, Tesoro de Patrimonio Histórico Andaluz, Tesoro de bienes muebles e inmuebles por destino de recintos religiosos, entre otros.

por ordenar nuestros términos basándonos en su frecuencia, esto es: primero el término más frecuente, luego el segundo más frecuente y así sucesivamente.

Considerando la forma en la que documentamos las terminologías del proyecto, la naturaleza de nuestras fuentes de información (catálogos/bases de datos relacionales) y el progreso en la documentación de nuestras colecciones no se puede aplicar ninguna de las técnicas tradicionales de estimación de frecuencias del léxico, como: la frecuencia relativa,²⁰ la frecuencia relativa por millón de palabras,²¹ el índice normalizado de dispersión,²² la frecuencia corregida, propuesta original de Lanke, según reporta Ham,²³ etcétera, ya que no es posible conocer el número total de ítems en las colecciones hasta que estas estén completamente catalogadas y, por ende, se desconoce el número total de términos que las describen.

Una alternativa plausible y aplicable a estos casos es estimar la frecuencia de documento, método que permite cuantificar el uso de una unidad léxica, forma o giro lingüístico a partir de su dispersión o recurrencia en fuentes de información, es decir, la frecuencia de documento de una unidad léxica es la frecuencia absoluta de la aparición de esta en diferentes fuentes de información.²⁴ Esto encuentra fundamento en lo explicado por Sinclair, de que solo los hechos léxicos más prominentes, que se repiten en una variedad de textos y autores de forma independiente deben considerarse para estar en una obra de referencia, principio que él llama *recurrence*.²⁵

Para aclarar el alcance de todo esto, pensemos en este ejemplo: el AFMT aporta al total de la muestra analizada 691 505 ocurrencias y 319 tipos; de estos, los cinco términos más frecuentes son “fotografía”, “transparencia”, “reprografía”, “diapositiva” e “imagen”; en oposición de las 10 347 ocurrencias y 72 tipos de la BD de la ASC, en los que los cinco términos más frecuentes son “dibujo”, “fotografía”, “estampa”, “postal” y “escultura”.

20 Charles Muller, *Estadística lingüística* (Madrid: Gredos, 1973), 27.

21 Douglas Biber, Susan Conrad, y Randi Reppen, *Corpus Linguistics: Investigating Language Structure and Use* (Cambridge: Cambridge University Press, 1998), 38.

22 Luis Fernando Lara Ramos y Roberto Ham Chande, “Base estadística del Diccionario del Español de México”, en *Investigaciones lingüísticas en lexicografía*, ed. Luis Fernando Lara Ramos, (Ciudad de México: El Colegio de México, 1974), 37-38.

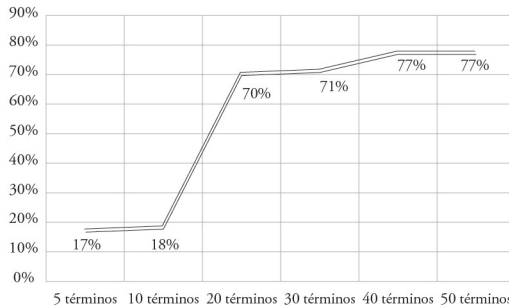
23 Lara Ramos y Ham Chande, 35-37.

24 Claudio Molina Salinas y Gerardo Eugenio Sierra Martínez, “Hacia una normalización de la frecuencia de los corpus crea y corde”, *Revista Signos. Estudios de lingüística* 48 (89), 2015, 318-324.

25 John Sinclair, “Corpora for lexicography”, en *A practical guide to lexicography*, ed. Piet Van Sterkenburg (Amsterdam/Philadelphia: John Benjamins Publishing Company, 2003), 167.

Si el primer término incluido y definido en el microtesauro hubiera sido “reprografía” y se usara para la catalogación en todos los museos y dependencias involucradas en el seminario, este término describiría 45 575 ítems del AFMT, CDA y UNIARTE, pero no le serviría a la ASC, el MUAC, el CIX o al MUCH, impidiendo, con esto, que el trabajo de documentación de las colecciones avance simultáneamente en todos los museos y dependencias.

Consecuentemente, aplicar la frecuencia de documento a nuestros listados terminológicos nos permitió establecer un ranquin en el que los primeros términos definidos servirían por igual a todos los participantes del seminario. En este caso, los cinco términos más frecuentes (por frecuencia de documento) que sirven para describir a los ítems de todas nuestras colecciones son “dibujo”, “escultura”, “pintura”, “fotografía” y “moneda”. Tan solo estos cinco términos describen al 17% de los registros de los catálogos de todas nuestras colecciones, es decir, 124 027 objetos. A continuación se puede ver gráficamente la cobertura que tienen los cinco, diez, veinte, treinta, cuarenta y cincuenta términos más frecuentes de nuestra lista respecto al total de ítems en nuestras colecciones.



3. Cobertura que tienen los términos, resultado de la aplicación de la frecuencia de documento.

Con base en la información de la gráfica anterior tomamos una decisión: comenzaríamos publicando en el microtesauro los cuarenta términos más frecuentes de nuestro listado, ya que con ello describimos dos terceras partes de los registros de nuestras BD, es decir, el 77% de registros capturados de nuestras colecciones o 552 102 ítems. Desde luego, estos cuarenta términos solo describen a los tipos de objetos artísticos y arqueológicos pero, más adelante, trabajaremos específicamente con tipos de materiales, técnicas y otras terminologías.

Estimar la frecuencia de documento de nuestros términos e integrarlos en una lista de frecuencias o ranquin de recurrencia también nos permite tomar decisiones sobre cuáles son los términos preferidos y cuáles los alternativos, esto con base en la posición que estos ocupen dentro de nuestra lista. De esta forma, tenemos evidencias cuantitativas de que “tarjeta postal” debe ser el término preferido por tener una frecuencia de documento mayor en nuestra lista, mientras que “postal” o “postcard” serían términos alternativos, ya que tienen una frecuencia menor.

En este momento no hemos identificado términos cancelados, pero suponemos que habría dos condiciones para reconocerlos, primero, una muy baja frecuencia en las BD; y, segundo, hacer un trabajo colegiado, como el que se describió al inicio de este apartado; aún falta explorar ambos procedimientos.

Por último, en esta etapa de trabajo es posible decidir si un término no fue debidamente normalizado; si es el caso, simplemente se regresa a la etapa anterior para corregir el procedimiento de normalización y se retoma el proceso.

2.3. Redacción de definiciones

Para redactar las definiciones es necesario contar con un corpus de referencia, precisar el tipo de definición que se va a aplicar y poner manos a la obra. Por una parte, en UNIARTE contamos con dos recursos documentales: el *Tesoro Terminológico de las Artes y Arqueología en México* (TTAAM), que es un corpus de definiciones en español que concentra 18 190 artículos terminográficos relativos a la historia del arte y arqueología;²⁶ y, por otro lado, el *Corpus de Documentos de la Historia del Arte y Arqueología en México* (CHAAM), que cuenta con 2 164 documentos relacionados con este dominio. Ambos recursos documentales son propiedad del IIE, están en línea y son de acceso restringido. En estos dos corpus se documentan los términos y, a partir de los contextos en los que aparecen, se infieren sus significados.

Sobre la definición, tomando en cuenta nuestras necesidades, hemos adoptado un tipo analítico o aristotélico en el que describimos la clase a la que pertenece nuestro definido (*genus*) y lo que lo cualifica, lo especifica o lo ca-

26 Claudio Molina Salinas, “Un Tesoro Terminológico de las Artes y Arqueología en México (TTAAM) desde el IIE-UNAM”, *Revista Electrónica Imágenes*, 2023.

racteriza (*differentia specifica*).²⁷ Esta diferencia específica siempre enuncia los materiales y técnicas que conforman al tipo de objeto, mientras que el *genus* es otro tipo de objeto más general (hiperónimo) o una superclase.

La estrategia planteada para facilitar la redacción de las definiciones, y que sirve para homogenizarlas, independientemente del redactor de estas, se basa en una propuesta basada en preguntas del tipo “completing”, planteamiento de Molina, Ruiz y Hernández.²⁸ A continuación se muestra la parte que retomamos de esta estrategia

La definición consta de dos partes:

1. *Genus* o género próximo. Puede ser una unidad léxica o una frase que identifique el tipo o la clase a la que esta unidad léxica pertenece. Para establecerlo es útil completar esta oración: “El (poner aquí al definido) es un tipo o clase de..._____”.

Ejemplos:

El cantor es un tipo o clase de cargo u. oficio.

El bajón es un tipo o clase de instrumento musical.

2. *Differentia* o diferencia específica. Califica o caracteriza a la unidad léxica definida; tiene la función de distinguirla de otros conceptos de la misma clase que se encuentran en relación horizontal. Generalmente se liga con el *genus* por medio de las formas léxicas: “que, cuyo, cual, entre otras, o, incluso, de cualquier participio”.

Ejemplo:

capellán de coro. Cargo u oficio *de quien* encabezaba el canto litúrgico a cambio de un pago fijo.

4. Modelo de preguntas que sirven para redactar definiciones analíticas (fragmento).

27 Robert Lew, “Identifying, Ordering and Defining Senses”, en *The Bloomsbury Companion to Lexicography*, ed. Howard Jackson (London/New Delhi/New York/Sydney: Bloomsbury Publishing, 2013), 296-97.

28 Claudio Molina Salinas, Antonio Ruiz Caballero, y Salvador Hernández Pech, “Un modelo de definiciones terminográficas para un glosario de documentos litúrgicos virreinales de México”, *Textos en Proceso* 6(1), 2020, 116.

Con base en todo lo anterior, comenzamos con la redacción de los borradores de las definiciones para, posteriormente, presentarlas a los especialistas del seminario.

2.4. Revisión de las definiciones y determinación de las relaciones jerárquicas

Los especialistas del seminario (dictaminadores de las definiciones en todos los casos) deben revisar las definiciones propuestas atendiendo a dos dimensiones: primero, preguntarse si el *genus* de la definición es, efectivamente, un tipo de objeto o una clase más generales que podrían recoger o incluir al definido; y, segundo, revisar cuidadosamente si los materiales y las técnicas que conforman la *differentia specifica*, en su conjunto, describen lo que particulariza o hace único al definido. El cotejo contra otras fuentes de autoridad no se descarta en este momento. Considerando estas dos dimensiones del significado, el dictaminador de la definición ya podría tener algunos elementos, adicionales a su experiencia profesional, para valorarla.

Naturalmente, la valoración y validación de las definiciones requiere un trabajo colegiado y una toma de decisiones de forma consensuada, por ello consideramos que el seminario permanente es un lugar conveniente para la discusión sobre la idoneidad de estas.

En esta etapa de trabajo es posible que se identifiquen problemas graves con la definición, por ello podría regresar a la etapa anterior y redactarse nuevamente; incluso, este análisis del significado puede hacer notar al analista problemas con la normalización del definido, así que también podría regresar el término a esa etapa de trabajo.

Por último, el *genus* de las definiciones determina las relaciones jerárquicas verticales en el microtesauro; de esta forma organizamos a “pintura al fresco” como un hipónimo de “pintura mural” y este, a su vez, como un hipónimo de “pintura”:

TOMATE: microtesauro de tipos de objetos, materiales y técnicas de manufactura que describen el patrimonio artístico y cultural de la UNAM

The screenshot shows the TOMATE microthesaurus interface. At the top, there is a navigation bar with 'Inicio', 'Menú', 'Agregar término', a search box with a 'Buscar' button, 'Búsqueda avanzada', and 'Sobre...'. Below the navigation bar is a horizontal menu with letters A, C, D, L, M, O, P, S, T. The main content area displays a hierarchical menu structure for 'Tomate: microtesauro' with the following items: '<Tomate: microtesauro>', '<Agentes>', '<Conceptos asociados>', '<Deterioros o amenazas>', '<Materiales>', '<Periodos espacio-temporales>', '<Técnicas>', '<Tipos de objetos>', '<Según los materiales y/o técnicas que los forman>', 'pintura', 'pintura mural', and 'pintura al fresco'.

5. Ejemplo de las jerarquías en el microtesauro.

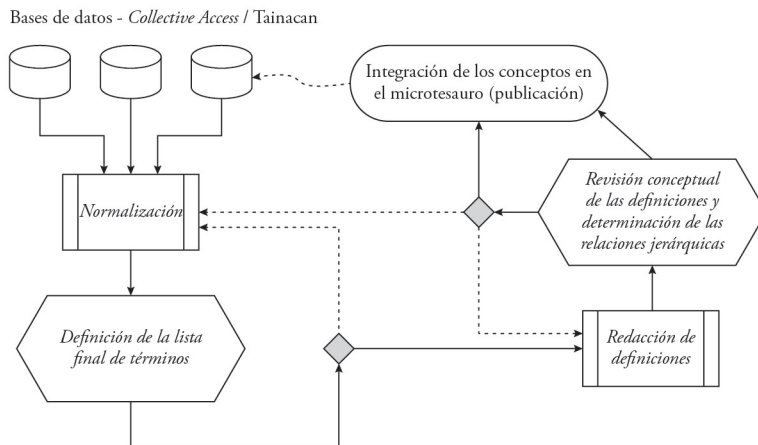
2.5. Integración de los conceptos en el microtesauro (publicación en TemaTres)

Para publicar el microtesauro hemos optado por utilizar el *software* TemaTres, que sirve para crear y editar todo tipo de vocabularios controlados: desde una lista de control terminológico simple hasta un tesaurus. TemaTres nos permite, entre otras, declarar relaciones entre términos (de equivalencia, jerárquicas y asociativas), redactar definiciones y distintos tipos de notas (de alcance, históricas, bibliográficas, etcétera), y administrar y gestionar distintos usuarios y editores; además, se puede documentar cada uno de los pasos explicados directamente en dicho *software* e ir publicando los avances paulatinamente.²⁹

2.6. Últimos ajustes a la propuesta metodológica

Después de poner en práctica de forma global esta metodología y hacer una evaluación de los cinco pasos originales de ella, descubrimos la necesidad de replantearla, sin afectar la definición de las etapas, pero implicando dos momentos de toma de decisiones al definir la lista final de términos y revisar las definiciones.

²⁹ Maria Antonia Osés, *Manual de TemaTres. La consulta y gestión de un vocabulario controlado* (Ciudad Autónoma de Buenos Aires, 2017), *passim*.



6. Propuesta metodológica para la creación de un microtesauro (diagrama de flujo, reelaborado).

Como se puede ver en el esquema anterior, esta toma de decisiones podría llevar a uno o varios casos trabajados tanto a la etapa de normalización como a la de redacción de definiciones (puntos señalados con los rombos).

3. RESUMEN Y CONCLUSIONES

En este capítulo hemos explicado el origen y contexto del trabajo de nuestro seminario, así como puesto en perspectiva la necesidad de contar con un microtesauro que complemente el uso de los *softwares* de administración, gestión y publicación de colecciones que usamos (*CollectiveAccess* y *Tainacan*).

Además, hemos propuesto una metodología de cinco pasos para publicar un microtesauro, explicada e ilustrada con ejemplos específicos que, a su vez, tiene dos momentos de toma de decisiones: justo cuando se definen la lista final de términos y se revisan las definiciones, respectivamente (véase antes: “6. Propuesta metodológica para la creación de un microtesauro (diagrama de flujo, reelaborado)”).

En nuestra opinión y experiencia, el proceso de documentación léxica y los tiempos de publicación del microtesauro se acortan notablemente con esta propuesta, misma que podría adoptarse en lugar de metodologías tradicionales y aplicarse a otros proyectos de catalogación semejantes a este.

Sabemos que incluir solo tipos de objetos es discutible y se podría cuestionar la representatividad del microtesauro en esta etapa, pero no hay que

perder de vista que el objetivo de este proyecto es representar a nuestras colecciones y no al dominio de la historia del arte y arqueología en su totalidad. Finalmente, como seminario, hemos argumentado que este procedimiento de compilación de las terminologías es un muestreo por selección intencionada o de conveniencia que, pese a ser subjetivo, sesgado y que es difícil determinar su representatividad³⁰ es suficiente para los alcances de nuestros proyectos individuales.

Por último, creemos que nuestro microtesauro debería tener un nombre más atractivo, que también describa sus alcances; por ello lo hemos llamado, tentativamente, *TOMATE: microtesauro de tipos de objetos, materiales y técnicas de manufactura que describen el patrimonio artístico y cultural de la UNAM*. TOMATE es el acrónimo de: tipos de objetos, materiales y técnicas.

30 Jordi Casal y Enric Mateu, “Tipos de Muestreo”, *Revista de Epidemiología y Medicina Preventiva* 1, 2003, 5.

4. REFERENCIAS BIBLIOGRÁFICAS

- Alcaraz Martínez, Rubén, “CollectiveAccess, un sistema de gestión y difusión de colecciones de museos, archivos y bibliotecas”, *BiD: textos universitarios de biblioteconomía i documentació* 33, 2014, <https://doi.org/10.1344/BiD2014.33.23>
- Baca, Murtha, Patricia Harpring, Elisa Lanzi, Linda Mcrae, y Ann Whiteside, eds., *Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images*. Chicago: American Library Association, 2006.
- Biber, Douglas, Susan Conrad y Randi Reppen, *Corpus Linguistics: Investigating Language Structure and Use*, Cambridge: Cambridge University Press, 1998.
- Casal, Jordi y Enric Mateu, “Tipos de Muestreo”, *Revista de Epidemiología y Medicina Preventiva* 1, 2003, 3-7.
- “CollectiveAccess”, [collectiveaccess.org](https://www.collectiveaccess.org/), 2022, Consultado el 30 de abril de 2024, <https://www.collectiveaccess.org/>
- Lara Ramos, Luis Fernando, *Curso de lexicología*, México, D.F.: El Colegio de México, 2006.
- Lara Ramos, Luis Fernando, *Teoría del diccionario monolingüe*, México, D.F.: El Colegio de México, 1997.
- Lara Ramos, Luis Fernando y Roberto Ham Chande, “Base estadística del Diccionario del Español de México”, en *Investigaciones lingüísticas en lexicografía*, editado por Luis Fernando Lara Ramos, 7-30, Ciudad de México: El Colegio de México, 1974.
- Lew, Robert, “Identifying, Ordering and Defining Senses”, en *The Bloomsbury Companion to Lexicography*, editado por Howard Jackson, 284-302, London/New Delhi/New York/Sydney: Bloomsbury Publishing, 2013.
- Ministerio de Educación, Cultura y Deporte, “Definición de Bien de Interés Cultural”, página del Ministerio de Educación, Cultura y Deporte, *s/f.*, consultado el 30 de abril de 2024, <https://www.cultura.gob.es/cultura/patrimonio/bienes-culturales-prottegidos/definicion.html>
- Molina Salinas, Claudio, “Un Tesoro terminológico de las artes y arqueología en México (TTAAM) desde el IIE-UNAM”, *Revista Electrónica Imágenes*, 2023, <http://www.revistaimagenes.esteticas.unam.mx/un-tesoro-terminologico>

- Molina Salinas, Claudio, Antonio Ruiz Caballero y Salvador Hernández Pech, “Un modelo de definiciones terminográficas para un glosario de documentos litúrgicos virreinales de México”, *Textos en Proceso* 6(1), 2020, 96-117.
- Molina Salinas, Claudio y Gerardo Eugenio Sierra Martínez, “Hacia una normalización de la frecuencia de los corpus CREA y CORDE”, *Revista Signos. Estudios de lingüística* 48, n.º 89, 2015, 307-31, <https://doi.org/10.4067/S0718-09342015000300002>
- Morales-del-Castillo, José Manuel, Pedro Ángeles Jiménez y Claudio Molina Salinas, “Mexico’s Tradition and Culture Entering the Digital Age: The Mexican Cultural Heritage Repository Project”, *Heritage* 2, 2019, 356-365, <https://doi.org/10.3390/heritage2010024>
- Moure Romanillo, Alfonso, *Patrimonio cultural y patrimonio natural: una reserva de futuro*, Santander: Universidad de Cantabria, 2003.
- Muller, Charles, *Estadística lingüística*, Madrid: Gredos, 1973.
- Osés, María Antonia, *Manual de TemaTres, La consulta y gestión de un vocabulario controlado*, Ciudad Autónoma de Buenos Aires, 2017, consultado el 30 de abril de 2024, <http://vocalarios.saij.gob.ar/portalthes/descargas/ManualTematresSAIJ2017.pdf>
- Salse, Marina, Javier Guallar-Delgado, Núria Jornet-Benito, María Pilar Mateo Bretos y Josep Oriol Silvestre-Canut, “GLAM metadata in museums and university collections: a state-of-the-art (Spain and other European countries)”, *Global Knowledge, Memory and Communication* 73(4/5), 2024, 477-95, <https://doi.org/10.1108/GKMC-06-2022-0133>
- Sinclair, John, “Corpora for lexicography”, en *A practical guide to lexicography*, editado por Piet Van Sterkenburg, 167-78, Amsterdam/Philadelphia: John Benjamins Publishing Company, 2003.
- “OpenRefine”, The Open Refine Web’s Page, 2012, consultado el 30 de abril de 2024, <https://openrefine.org/>.
- “Tainacan Wiki”, The Tainacan Project Web’s page, Consultado el 30 de abril de 2024, <https://tainacan.github.io/tainacan-wiki/#/?id=tainacan-wiki>

Unidad de Información para las Artes (UNIARTE), “Unidad de Información para las Artes | Instituto de Investigaciones Estéticas”, Instituto de Investigaciones Estéticas, 2020, consultado el 30 de abril de 2024, <http://www.esteticas.unam.mx/uniarte>

Visual Resources Association, “VRA Core 4.0 Element Description”, Visual Resources, 2007, consultado el 30 de abril de 2024, https://www.loc.gov/standards/vracore/vra_Core4_Element_Description.pdf

Organización del conocimiento. Bibliotecología, Terminología e Información. Instituto de Investigaciones Bibliotecológicas y de la Información/UNAM. Edición digital. Coordinación editorial: Sergio Javier Sepúlveda H. y Angélica Valenzuela; revisión especializada: Diana Martínez Aguilar; corrección de pruebas: Carlos Ceballos Sosa y Marcos Emilio Bustos Flores; formación editorial y portada: Carlos Papaqui Landeros. Se terminó en febrero de 2025.