



Ciencias de la Información
Facultad: Ciencias Económicas
Trabajo de Diploma

Para optar por el Título de
**Licenciado en Ciencias de la
Información**

*Desarrollo de un Tesauro especializado en el área de la
Ciencia de los Materiales para el Repositorio Institucional
Nínive de la Universidad de Moa Dr. “Antonio Núñez
Jiménez”.*

Ydelsi Vielza Caraballo

Moa, 2019



Ciencias de la Información
Facultad: Ciencias Económicas

Trabajo de Diploma

Para optar por el Título de
Licenciado en Ciencias de la
Información

Desarrollo de un Tesauro especializado en el área de la Ciencia de los Materiales para el Repositorio Institucional Nínive de la Universidad de Moa Dr. “Antonio Núñez Jiménez”.

Autora: Ydelsi Vielza Caraballo

Tutores: Dr. C. José Luis Montero O' farrill

Lic. Niurka de la Vara Garrido

Moa, 2019

Declaración de autoría

Declaro que soy la única autora de este trabajo y autorizo al Departamento de Ciencias de la Información de la Universidad de Moa Dr. Antonio Núñez Jiménez para que hagan el uso que estimen pertinente con este trabajo.

Para que así conste firmo la presente a los _____ días del mes de _____ del _____.

Ydelsi Vielza Caraballo

Firma de la autora

Lic. Niurka de la Vara Garrido

Firma del tutor

Dr. José Luis Montero O´farill

Firma del tutor

Pensamiento

*“No hace falta ser un Sol, basta con ser una pequeña voz,
pero que alumbre allí, donde hace falta”.*

Simón Bolívar

Dedicatoria

A:

Mis padres Ydalia e Idel Esteban quienes con su amor, paciencia y esfuerzo me han permitido llegar a cumplir hoy un sueño más, gracias por inculcar en mí el ejemplo de esfuerzo y valentía, de no temer a las adversidades para hacer de mí una mejor persona. Gracias por sus consejos por sus oraciones, consejos y palabras de aliento.

A mis hermanos Albero y Osvaldo por su cariño y apoyo incondicional, durante todo este proceso, por estar conmigo en todo momento gracias.

A toda mi familia porque con sus oraciones, consejos y palabras de aliento hicieron de mí una mejor persona y de una u otra forma me acompañan en todos mis sueños y metas.

A mi abuela Virginia, que ya no está a mi lado, pero su cariño prevalece siempre en mi corazón, fue una mujer tolerante, honesta, bondadosa y generosa, siendo una fuente de inspiración en cada paso de mi vida.

Mis sobrinas Melody Charlotte y Nathaly, porque a pesar de la distancia, el ánimo y alegrías que me brindan me da la fortaleza para seguir adelante, para que vean en mí un ejemplo a seguir.

Agradecimientos

Gracias a mis padres por ser los principales promotores de mis sueños, gracias por cada día confiar y creer en mí y en mis expectativas. Gracias a mi madre por su fe, generosidad y su incansable ayuda en todo momento, gracias a ella he llegado a culminar un peddño más en mi vida. Gracias a mi padre por siempre desear y anhelar lo mejor para mí, por su apoyo ilimitado.

A mis hermanos: Alberto y Osvaldo, por formar parte de mí, por brindarme su cariño, su amor y darme la felicidad de contar con ellos.

A mis familiares de Moa en especial a Maryolis, Margiolis y Maira por ser estupendas mujeres, amigas y hermanas las cuales me apoyaron durante mis estudios, gracias por su cariño, su amor sincero y ser especiales.

A toda mi familia, por llenarme de amor, por ayudarme a formar mis valores, mis principios, mi perseverancia y mi empeño, por darme fuerzas y ser una lumbrera que me colma de ímpetu cada día.

A Dios, por darme la oportunidad de vivir y por estar conmigo en cada paso que doy, por fortalecer mi corazón e iluminar mi mente y por haber

*puesto en mi camino a aquellas personas que han sido mi soporte y
compañía durante todo el periodo de estudio.*

*A mi novio Eudis por su amor, paciencia, apoyo incondicional y que ha
sido mi amigo en todo momento y a su familia por ser personas
maravillosas. Gracias por amarme como solo tú lo sabes hacer.*

*A todos mis compañeros de aula, que más que amigos han sido una
familia en estos 5 años en momentos de sacrificio, entrega, desvelos y
alegrías, en especial a Luis, Anita, Liz, Yixian, Elianis, Mailén Yeilenis,
Yanaila, Anniaa todos en general.*

*A mis amistades en especial a Aylín, Niubel, Meibis, Luismar y familia,
por apoyarme, por su generosidad, su comprensión y su ayuda
desinteresada en momentos difíciles; por buenos momentos de felicidad y
alegría.*

*A mi tutora Niurka, quien desde el primer momento me brindó su
amistad y su bondad. Por sus útiles consejos, por impulsarme hacia la
victoria y guiarme hacia este feliz término.*

*A mi tutor José Luis, por su orientación, valiosos comentarios,
sugerencias y conocimientos brindados.*

*Al claustro de profesores que durante estos 5 años me formaron y
aportaron los conocimientos esenciales para obtener buenos resultados.*

*Por compartir conmigo sus conocimientos, por enseñarme a ser útil, a
emprender metas y a luchar.*

*A quienes jamás encontraré la forma de agradecer el cariño, comprensión
y apoyo brindado en los momentos buenos y malos de mi vida, hago este
triunfo compartido y a todos los que de una forma u otra contribuyeron
en la realización de la tesis.*

A todos GRACIAS

.

.

RESUMEN

El uso de vocabularios controlados constituye una alternativa para la representación normalizada de los términos y el posterior acceso a la información. Con esta investigación se desarrolla un tesoro especializado en el área de la Ciencia de los Materiales para el Repositorio Institucional Nínive de la Universidad de Moa Dr. Antonio Núñez Jiménez. Para ello se utilizó el Sistema de Clasificación Decimal Dewey y los planes de estudio de las especialidades de la Facultad de Metalurgia-Electromecánica, así como varias herramientas léxicas que permitieron precisar la terminología. En su planeamiento se empleó un procedimiento mixto utilizando el método de comité y el método empírico, además del uso del software TemaTres. Como resultado se obtuvo un tesoro especializado con un total de 1603 términos normalizados para ser utilizados en el proceso de indización y recuperación de la información en el Repositorio. El tesoro de Ciencia de los Materiales refleja las expresiones utilizadas por los autores de los documentos y los usuarios en sus formulaciones de búsqueda de información.

Palabras clave: tesauros, vocabularios controlados, indización, ciencia de los materiales.

ABSTRACT

The use of controlled vocabularies constitutes an alternative for the normalized representation of terms and the subsequent access to information. With this research, a thesaurus specialized in the area of Materials Science is developed for the Nínive Institutional Repository of the University of Moa Dr. Antonio Núñez Jiménez. Due to, the Dewey Decimal Classification System and the curricula of the specialties of the Faculty of Metallurgy-Electromechanics were used, as well as several lexical tools that allowed to specify the terminology. In their planning, a mixed procedure was used using the committee method and the empirical method, in addition to the use of the TemaTres software. As a result, a specialized thesaurus with a total of 1603 standardized terms was obtained to be used in the process of indexing and retrieving the information in the Repository. The Materials Science thesaurus reflects the expressions used by the authors of the documents and the users in their information search formulations.

Key words: thesauri, controlled vocabularies, indexing, material science.

ÍNDICE

INTRODUCCIÓN	1
CAPÍTULO I: ASPECTOS TEÓRICOS-CONCEPTUALES DEL PROCESO DE ELABORACIÓN DE TESAUROS.....	7
I.1. La indización. Conceptos y generalidades.....	7
I.1.1. Atributos esenciales en el proceso de indización.	8
I.2. Conceptualización y características generales de los lenguajes documentales. 9	
I.2.1. Tipología de lenguajes documentales.....	9
I.2.2. Objetivos y funciones del lenguaje documental.	11
I.3. Los vocabularios controlados.....	13
I.3.1. Vocabularios controlados como formas para organizar el conocimiento. ..	13
I.3.2. Lenguajes controlados.....	14
I.3.3. Tipos de vocabularios controlados (Albuquerque, 2018).....	17
I.3.4. El control del vocabulario en el entorno digital.....	19
I.3.5. Los vocabularios controlados y su uso en los repositorios.	20
I.4. Herramientas de indización.....	22
I.5. Los Tesauros	22
I.5.2. Objetivos del tesoro.	24
I.5.3. Normas y directrices para la elaboración de tesauros.	25
I.5.6. Existen algunas condiciones que por lo general debe cumplir todo tesoro:	26
I.5.7. Estructura de los tesauros.	27
I.5.8. Métodos para la construcción de tesauros.	30
CAPÍTULO II: METODOLOGÍA UTILIZADA PARA LA ELABORACIÓN DEL TESAURO ESPECIALIZADO EN EL ÁREA DE LA CIENCIA DE LOS MATERIALES.	34
II.1. Caracterización general del Repositorio Nínive.	34
II.2 Alcance temático del tesoro.	34
II.3 Población objeto de estudio.	35
II.4. Guía metodológica de trabajo para la confección del tesoro.	35
CONCLUSIONES.....	61

RECOMENDACIONES	62
REFERENCIA BIBLIOGRÁFICA.....	63

INTRODUCCIÓN

En la búsqueda de información, el éxito o el fracaso del resultado obtenido dependerán, en gran medida, de la correcta elección de los términos por los cuales el contenido de los documentos ha sido identificado. En este sentido los documentos que componen cualquier unidad significativa de información almacenada, deben ser analizados de manera tal de que sea posible su recuperación.

El tema, los conceptos que implícita o explícitamente están recogidos en una obra se extraen mediante el “Análisis de Contenido”, cuyas operaciones, como la indización, dan al lector un conocimiento más o menos profundo de la información que contiene. La operación de indizar consiste, en el análisis e identificación de los conceptos del documento y la selección de aquellas nociones que representen con mayor fidelidad la información que contiene. Para normalizar la denominación de dichos conceptos se procede a su traducción a un lenguaje documental, lo cual facilita la recuperación, independientemente del analista que trate el documento. (Cañedo Andalia y Small Chapman, 2011)

Por eso la recuperación de la información es un tema que cobra gran importancia con el crecimiento cada vez mayor de los volúmenes de información y el desarrollo de las Tecnologías de la Información y las Comunicaciones. Podemos afirmar que: el éxito de la mediación documental está sujeto al empleo de estrategias de búsqueda con lenguajes controlados, garantizando que un único término denomine un único objeto, y un objeto no pueda ser definido por más de un término.

En los últimos años los Repositorios Institucionales (RI) han cobrado importancia en la sociedad académica y científica, porque representan una fuente de información digital especializada, organizada y accesible para los lectores de diversas áreas. Los RI son sistemas informáticos dedicados a gestionar los trabajos científicos y académicos de diversas instituciones de forma libre y gratuita, es decir, siguiendo las premisas del movimiento *Open Access* (OA).

Actualmente en Cuba, comienza cierto auge del desarrollo e implementación de los repositorios institucionales. En particular, los Centros de Educación Superior (CES) basados en la “Política para el desarrollo del Sistema de Repositorios Digitales en la

Red de Bibliotecas Universitarias del Ministerio de Educación Superior” van teniendo un buen desarrollo. (Machado Rivero, 2017)

En este escenario, donde todos los CES deben tener un Repositorio Institucional con el objetivo de convertir los conocimientos científicos en instrumentos a disposición del avance y el desarrollo a nivel nacional e internacional, cobra utilidad establecer modelos de indización como instrumentos léxicos que permitan tanto a usuarios como a profesionales de la información organizar y acceder de manera rápida y precisa al conocimiento científico que en ellos se almacena.

Es de esperar, por lo tanto, que se implementen herramientas que permitan la más eficaz y eficiente recuperación de información, en los RI. Estos sistemas deben contemplar el uso de vocabularios controlados para normalizar la información de los recursos, mejorando aspectos como la carga de metadatos, la interoperabilidad entre sistemas y la búsqueda y difusión de contenido. (Bernal, 2016)

El uso de los vocabularios controlados en los RI favorece múltiples procesos ya que define un conjunto limitado de términos para referirse a un único concepto de manera consistente unificando los encabezamientos o puntos de acceso de un catálogo bibliográfico.

En el Repositorio Institucional Nínive de la Universidad de Moa Dr. Antonio Núñez Jiménez con el incremento de la información depositada en el dominio temático de la Ciencia de los Materiales y al no desarrollarse un control del vocabulario durante el procesamiento de esta información, se ha generado un índice de materias muy amplio y diverso lo que dificulta la consistencia en los metadatos de materia y la búsqueda y recuperación de la información más precisa y efectiva.

En tal sentido en la búsqueda de información efectuada para esta investigación se confirma la existencia de varios tesauros tales como el de la UNESCO, EUROVOC, DOCUTES y el tesoro de Geología los cuales en alguna medida han sido utilizados en el proceso de indización de los documentos alojados en el repositorio, sin embargo en el dominio temático de la Ciencia de los Materiales es donde se ha confrontado mayores dificultades ya que estos tesauros no se ajustan en lo fundamental a las expresiones utilizadas por la comunidad científica del Repositorio Institucional Nínive en dicha área temática.

Por lo que se plantea como **problema científico** de esta investigación:

¿Cómo desarrollar un Tesauro especializado en el área de la Ciencia de los Materiales para el Repositorio Institucional Nínive de la Universidad de Moa Dr. Antonio Núñez Jiménez?

Estableciendo como **objeto de estudio**: los vocabularios controlados.

Campo de acción: Tesauro especializado en el área de la Ciencia de los Materiales del Repositorio Institucional Nínive.

Objetivo general:

- Desarrollar el Tesauro en el dominio temático de la Ciencia de los Materiales, para el Repositorio Institucional Nínive de la Universidad de Moa Dr. Antonio Núñez Jiménez.

Objetivos específicos:

- Determinar los aspectos teóricos-conceptuales sobre los tesauros, particularizando su uso en los repositorios institucionales.
- Establecer la metodología y lineamientos a utilizar para el desarrollo del Tesauro especializado en el área temática de la Ciencia de los Materiales.
- Identificar las herramientas léxicas en el campo de la Ciencia de los Materiales para extraer los términos que se adaptan al contexto de la colección documental seleccionada.
- Establecer la normalización de los términos temáticos en el área de la Ciencia de los Materiales para el Repositorio Institucional Nínive de la Universidad de Moa Dr. Antonio Núñez Jiménez.
- Generar el tesauro especializado en la Ciencia de los Materiales haciendo uso del software Tema Tres.

Idea a Defender:

La implementación de un Tesauro especializado en el dominio temático de la Ciencia de los Materiales favorecerá los procesos de indización, búsqueda y recuperación de

la información de los recursos digitales disponibles en el Repositorio Institucional Nínive de la Universidad de Moa Dr. Antonio Núñez Jiménez.

Para el logro de estos objetivos se plantean las siguientes tareas:

- Determinar los aspectos teóricos conceptuales sobre el control de los vocabularios en los RI.
- Determinar la metodología para desarrollar tesauros.
- Establecer los lineamientos para desarrollar tesauros.
- Seleccionar las fuentes terminológicas a utilizar para precisar los términos candidatos a incorporar al tesoro
- Compilar de todos los términos del dominio temático de la Ciencia de los Materiales utilizados en el proceso de indización de los documentos representados en el Repositorio Institucional Nínive de la Universidad de MoaDr. Antonio Núñez Jiménez.
- Someter a criterio de expertos la terminología compilada.
- Generación del tesoro.
- Divulgación del producto terminado.

Métodos de la investigación:

Se utilizó el **Materialismo Dialéctico e Histórico** como método general de investigación para analizar los fenómenos históricos y sociales que dieron lugar al lenguaje documental. Se explican los conceptos, métodos y herramientas que posibilitaron el desarrollo del vocabulario a partir del avance tecnológico facilitando el proceso de búsqueda debido al cúmulo de recursos informativos.

Métodos teóricos:

- **Análisis y Síntesis:** Posibilitó la división del problema en sus partes y la sistematización de los referentes teóricos que sustentan el control de términos temáticos. Además permitió el procesamiento e interpretación de los

resultados del diagnóstico realizado para constatar el estado actual del problema científico.

- **Deducción e inducción:** Permitió el análisis a partir de los métodos, técnicas y herramientas establecer relaciones a partir de la deducción y el razonamiento para permitir la validez de los datos y luego ser aplicados de forma eficaz.
- **Sistémico:** Se utilizó para determinar sistémicamente los rasgos fundamentales del control temático y profundizar en los elementos esenciales integrando los conocimientos para una mejor interpretación.

Métodos empíricos

- **Análisis Documental:** Se realizó un estudio bibliográfico exhaustivo en fuentes de información que abordan la temática, lo que permitió profundizar en los aspectos teóricos conceptuales y metodológicos para la elaboración de tesauros.
- **Criterio de expertos:** Se aplicó esta técnica a los investigadores y expertos en el campo de la Ciencia de los Materiales específicamente en los ejes temáticos de la Mecánica, Eléctrica, Química y Metalurgia, para validar el núcleo del vocabulario.
- **Observación participante:** Esta técnica se utilizó para observar el comportamiento de los usuarios en las búsquedas de información y el modo de utilizar el repositorio durante el uso de las colecciones en la materia de la Ciencia de los Materiales, participando en la conducción de las búsquedas de información realizadas por los usuarios.
- **Entrevista:** Se empleó para conocer el criterio de investigadores y expertos en el tema a abordar.

Estructura capitular:

El contenido de este trabajo de tesis consta de 2 capítulos, conclusiones, recomendaciones y bibliografía y anexos.

El primer capítulo hace referencia a los aspectos teóricos, metodológicos y conceptuales del proceso de indización con el uso de vocabularios controlados, especificando en el uso de estas herramientas en los RI. Se brinda además, una breve caracterización de algunos tesauros existentes utilizados en el proceso de indización para el repositorio. Se presentan experiencias aplicadas en contextos universitarios nacionales e internacionales.

En el segundo capítulo se describe todo el proceso seguido para el desarrollo del Tesauro especializado en el área temática de la Ciencia de los Materiales.

Se brindan conclusiones, recomendaciones y la bibliografía consultada, así como anexos.

CAPÍTULO I: ASPECTOS TEÓRICOS-CONCEPTUALES DEL PROCESO DE ELABORACIÓN DE TESAUROS.

La principal función de un sistema de información es poner a disposición de los usuarios la información relevante a sus necesidades para lo cual tiene que realizar diferentes procesos documentales, de lo general a lo particular: **indización, lenguajes documentales, vocabularios controlados** y el que ocupa esta investigación, **el tesauro**.

Para llegar a la creación del tesauro se debe comprender el proceso de indización para el establecimiento de un lenguaje documental que conduce a un vocabulario controlado.

I.1. La indización. Conceptos y generalidades.

“La indización es una de las etapas del procesamiento analítico sintético de la información (organización), por medio del cual se representa en algún portador material características esenciales de los documentos que permiten su posterior recuperación sin tener que revisar toda la colección. Esta busca la identificación y asignación de rubros, descriptores o encabezamientos de materia representativos, en esencia, de los contenidos principales y secundarios, tratados en un documento con el objetivo de recuperar información sobre un tema determinado”. (Cañedo Andalia y Small Chapman, 2011, p.)

El proceso de indización de un documento es parte de un mecanismo de recuperación de un sistema de información y determina e influye de modo importante, en la habilidad del sistema para responder a las solicitudes de los usuarios.

Precisamente, por su ostensible dualidad, se estudia de dos maneras (Cañedo Andalia y Small Chapman, 2011):

- Como proceso: consiste en representar el contenido y las características de un documento con la ayuda de los conceptos y aspectos considerados en un vocabulario controlado.
- Como finalidad: permite la búsqueda eficaz de la información contenida en un fondo documental. La indización es un requisito importante para una adecuada recuperación de la información.

La indización comprende dos fases de trabajo fundamentales (Cañedo Andalia y Small Chapman, 2011):

1. En la primera se somete el contenido del documento en lenguaje natural a un análisis conceptual, o sea la identificación de los contenidos tratados en el documento objeto de interés.
2. La conversión de las palabras y frases utilizadas en el lenguaje natural para representar una realidad en términos de un lenguaje controlado que pretende elevar la calidad de la recuperación de la información. Se representan los contenidos claves extraídos en el lenguaje del sistema. Se representan los contenidos claves extraídos en el lenguaje del sistema.

Entre los factores determinantes del éxito en la recuperación de información se encuentra el dominio del usuario del vocabulario empleado para representar los documentos de interés y las solicitudes de búsqueda.

I.1.1. Atributos esenciales en el proceso de indización.

Según Cañedo Andalia y Small Chapman (2011) en la indización, la especificidad y la exhaustividad constituyen los dos atributos cardinales. Estos son esenciales para determinar la clase de recuperación que se realizará.

La exhaustividad se define por el grado de cobertura que presentan los tópicos o asuntos tratados en los documentos o solicitudes en los resultados de la indización, es decir, en el producto. Cuanto mayor sea la representación de los tópicos expuestos en un documento o solicitud en su indización, mayor será la exhaustividad del proceso.

Se expresa cuantitativamente por el número de términos utilizados para realizar la representación. Comprende tópicos centrales y colaterales (principales y secundarios).

Por su parte, la especificidad es el nivel de detalle y exactitud con que se representa cada tópico tratado en un documento o solicitud en la indización.

La recuperación de información de manera eficiente requiere de la determinación de los términos más específicos que representan los conceptos presentes en un documento o solicitud. La representación de un concepto particular mediante un término, cuyo significado es más general que el concepto tratado en la entidad que

se indiza, conduce a una pérdida de especificidad y, por consiguiente, de información.

I.2. Conceptualización y características generales de los lenguajes documentales.

Según Gil Urdicián (1996 como se citó en Martín Gavilán, 2011) entiende por lenguaje documental "todo sistema artificial de signos normalizados, que facilitan la representación formalizada del contenido de los documentos para permitir la recuperación, manual o automática, de información solicitada por los usuarios". (p.1)

En la bibliografía existen muchas más definiciones; pero esta engloba las características fundamentales para esta investigación.

I.2.1. Tipología de lenguajes documentales

➤ Dependiendo del grado de control terminológico, se pueden distinguir (*Martín Gavilán, 2009*):

❖ **Lenguajes libres:** Lenguajes fundados en el principio de postcoordinación, se componen de términos no predefinidos que se va generando a partir de la realización de procesos de indización. Lenguajes altamente pertinentes, coherentes, de bajo coste y que proporcionan una enorme cantidad de resultados cuando se realizan búsquedas. De este tipo son las listas de descriptores libres y las listas de palabras clave (Uniterm, Índices permutados, Precis...). Los lenguajes libres no son propiamente lenguajes documentales puesto que para que reciban este nombre el vocabulario ha de estar controlado.

❖ **Lenguajes controlados:** Son lenguajes controlados los lenguajes documentales propiamente dichos: tesauros, listas de encabezamientos de materia y clasificaciones. Presentan un vocabulario previamente elaborado, y admiten un limitado número de modificaciones en el momento de su utilización. El vocabulario controlado proporciona al usuario un punto de búsqueda, en vez de dos o más, y reduce la posibilidad de que la búsqueda sea incompleta. Sin embargo, puede perderse alguna información debido a la falta de especificidad y errores en el análisis.

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

- ❖ **Lenguajes precoordinados:** Combinan los términos en el momento de la descripción. Son lenguajes precoordinados las clasificaciones y las listas de encabezamientos de materia. En estos sistemas las distintas nociones o conceptos que se unen para expresar una materia o un tema ocupan un lugar determinado, es decir se introducen en el momento de la indización en un orden previamente establecido y la recuperación habrá de hacerse secuencialmente, siguiendo ese orden. Suelen ser muy precisos, pero también muy rígidos.
- ❖ **Lenguajes postcoordinados:** Combinan los términos en el momento de la recuperación. Son lenguajes postcoordinados los tesauros (lenguaje utilizado para la indización por descriptores), y si incluyéramos los lenguajes libres podemos mencionar también las listas de descriptores libres y las listas de palabras clave. En estos sistemas las nociones o conceptos que se extraen en la indización para expresar el tema o los temas del documento tienen todas las mismas categorías y no se expresarán en ningún orden determinado. Serán los intereses de los usuarios en la fase de recuperación los que obliguen a la combinación de conceptos que se realizará utilizando los operadores booleanos, los operadores de comparación, etc.
 - Dependiendo de su estructura, se pueden distinguir:
- ❖ **Lenguajes de estructura jerárquica:** A la estructura jerárquica, también denominada arbórea o sistemática, responden determinadas clasificaciones (las clasificaciones jerárquicas): por ejemplo, la Dewey y la Clasificación Decimal Universal (CDU). En ellas los conceptos se distribuyen en clases o categorías, donde dependen unos de otros. Estos lenguajes van de lo general a lo particular. Útiles en fondos enciclopédicos, resultan poco operativos para recuperar información, son sistemas muy rígidos y difíciles de actualizar, pueden servir para organizar colecciones.
- ❖ **Lenguajes de estructura asociativa o combinatoria:** Son aquellos que permiten la libre combinación de términos entre sí, eliminando la rigidez de los lenguajes de estructura jerárquica. A la estructura combinatoria o asociativa responden, por ejemplo, las listas de encabezamientos de materia y los

tesauros, si bien es cierto que en la estructura de estos últimos participa también la estructura jerárquica. Estos lenguajes tienen una presentación alfabética y los conceptos se asocian o combinan libremente entre sí para expresar el tema o los temas.

- ❖ **Lenguajes de estructura sintáctica:** Surgen por la cantidad de ruido que producen los lenguajes combinatorios, que pueden poseer una red de relaciones semánticas, pero no tienen gramática. La estructura sintáctica permite concretar y ordenar los conceptos. Este tipo de lenguaje requiere un cierto grado de automatización en los procesos tanto de indización como de recuperación, por lo que su aparición está asociada al desarrollo tecnológico. Según su complejidad, se pueden diferenciar dos tipos: lenguajes sintácticos de gramática simple y lenguajes sintácticos de gramática elaborada.

Se consideran lenguajes documentales:

- a) Los **lenguajes de indización**, denominados también lenguajes de estructura combinatoria, que permiten representar el contenido de los documentos y de las consultas de forma analítica; son ejemplos de este tipo los tesauros, los léxicos, los índices.

Los lenguajes de indización comprenden los lenguajes libres, basados en la indización en lenguaje natural de los documentos, y expresados por medio de las listas de palabras clave y de descriptores libres; y los lenguajes controlados, basados en la representación unívoca de conceptos y términos, y expresados principalmente mediante el tesoro.

- b) Los **lenguajes de clasificación o de estructura jerárquica**, utilizados para representar el contenido de forma sintética; aquí podemos citar las clasificaciones, numéricas o jerárquicas, las listas de encabezamientos de materia.

I.2.2. Objetivos y funciones del lenguaje documental.

El lenguaje documental interviene en dos fases del proceso documental, en el momento de la descripción y en el de la recuperación de la información. El objetivo de dichas operaciones es el de facilitar la recuperación de la documentación reduciendo el esfuerzo y gasto de tiempo del usuario.

El lenguaje documental tiene capacidad para **representar** los mensajes contenidos en los documentos, lo que le permite cumplir dos objetivos fundamentales, el de **normalización** mediante un léxico apropiado capaz de reflejar con precisión la esencia conceptual de los documentos y el de **inducción** para establecer las referencias cruzadas entre sus términos, estando encaminadas a este último todas las demás funciones que desempeña a lo largo del proceso documental.

Además, reduce considerablemente el volumen de términos del lenguaje natural no tomando en consideración más que los sustantivos o los sintagmas nominales, seleccionando además, de entre ellos, un solo término entre dos o más nociones consideradas sinónimos (Almazán, 2012). En virtud de estas características, se le reconoce un carácter mediador en los procesos de resumen e indización al controlar la posible ambigüedad entre términos.

En la fase de descripción, tras la lectura e identificación de los conceptos contenidos en los documentos, se representan mediante un lenguaje documental, que proporciona un vocabulario unívoco que favorece la coherencia del análisis y permite traducir los conceptos en términos normalizados.

Pero el objetivo más importante sería conseguir la consistencia en la indización:

“será más probable que los indizadores estén de acuerdo sobre los términos necesarios para representar determinada materia, si éstos son seleccionados de una lista previa, que si pueden utilizar cualquier término. Lo mismo ocurre en el proceso de búsqueda: será más fácil identificar los términos apropiados a una necesidad de información si son seleccionados de una lista determinada. Por lo tanto, el vocabulario controlado facilita la coincidencia entre el lenguaje de los indizadores y el de los usuarios que hacen la búsqueda”. (Lancaster, 2002 como se citó en Soler Monreal, 2009)

Según Van Slype(1991 como se citó en Vizcaya Alonso, 1999) considera que los lenguajes de indización pueden intervenir en seis momentos diferentes del proceso de búsqueda:

1. Selección de los sistemas documentales que se van a interrogar.
2. Enunciado de los conceptos de la pregunta, en lenguaje natural.
3. Traducción a un lenguaje de indización.

4. Formulación de la ecuación.
5. Extensión asistida por el ordenador.
6. Apreciación final de la pertinencia.

Además de las mencionadas funciones, el lenguaje documental es de gran utilidad para la ordenación o archivo de documentos. Resuelve igualmente problemas planteados por el multilingüismo, como en el caso de los tesauros multilingües, que sirven de puente a diferentes lenguas, o en el caso de los lenguajes de clasificación, fundados en la simbolización numérica, que los hace independientes de toda lengua y constituyen una escritura comprensible en todos los idiomas, de ahí su fácil empleo internacional.

I.3. Los vocabularios controlados

El ser humano comenzó hace miles de años a generar, representar y transferir conocimiento. Los recursos usados en la representación del conocimiento, a saber: recursos auditivos, recursos visuales y recursos lingüísticos; para una adecuada organización y acceso a dichos recursos se han ido construyendo distintas herramientas, basadas en mayor o menor medida en representaciones conceptuales reducidas directamente de la representación del conocimiento plasmada en los documentos. Ejemplos de instrumentos para la organización y accesibilidad del conocimiento son los *vocabularios controlados*. (Cámara de la Fuente, 2004 como se citó en Soler Monreal y Gil Leiva, 2010)

I.3.1. Vocabularios controlados como formas para organizar el conocimiento.

¿Desde cuándo se habla de control de vocabulario en el ámbito de la Organización del Conocimiento?

Según Barité (2014), no hay para esta pregunta una respuesta precisa y ajustada. No obstante, es posible aseverar que casi seguramente el término fue extrapolado desde el área del vocabulario básico para el aprendizaje y el estudio de las lenguas hacia la Bibliotecología y la Ciencia de la Información.

En parte como reacción al Basic English surgió el Movimiento del Vocabulario Controlado (*Vocabulary Control Movement*), el cual de la mano de Harold Palmer y Michael West generó alternativas fundamentales y precursoras de los diccionarios de aprendizaje de una lengua y de las listas de palabras altamente estructuradas. Este

movimiento aportó un nuevo método que “intentó buscar criterios válidos para la selección de vocabulario con la finalidad de simplificar los textos utilizados en la enseñanza de la lectura en lengua extranjera” (Esteban García, 2007 como se citó en Barité, 2014), dando primacía al criterio de frecuencia de las palabras en la lengua.

En forma paralela, hubo ya en la literatura de la Bibliotecología y la Ciencia de la Información en la década de los 40, varias aproximaciones a la utilización de la palabra “control” en forma aislada, pero expresiva.

En lo que hace a la Organización del Conocimiento, la palabra “control” siguió siendo utilizada en los últimos cincuenta años en dos direcciones principales: (a) por un lado, vinculada a los procesos de normalización de la forma gráfica de los términos de indización y de las relaciones establecidas entre ellos, genéricamente denominados como procesos u operaciones de “control de vocabulario”; (b) por otro lado, asociada a la identificación de los repertorios surgidos de esos procesos, que comenzaron a denominarse “listas controladas”, “índices controlados”, “lenguajes controlados” y finalmente “vocabularios controlados”, expresión esta última, que parece haberse establecido de modo definitivo en la terminología del área, sin perjuicio de la aparición esporádica de las otras expresiones.

I.3.2. Lenguajes controlados

El propio término «lenguaje controlado» se utiliza en diferentes contextos con diferentes acepciones. Por ejemplo, en el ámbito de la documentación y las ciencias de la información, el término lenguaje controlado suele emplearse como sinónimo de vocabulario controlado, un tipo de lenguaje documental utilizado para indizar y recuperar información de documentos. (Laguens García, 2006 como se citó en Ramírez Polo, 2012)

Sin embargo, otros autores entienden que los lenguajes controlados son variantes del lenguaje normativo o estándar (Nyberg, Mitamura,yHujisen, 2003; Lehrndorfer, 1996; Hujisen, 1998 como se citó en Ramírez Polo, 2012).

Según Ramírez Polo (2012) el lenguaje controlado pretende controlar el vocabulario y las estructuras sintácticas de un sublenguaje, pero además incluye vocabulario y estructuras sintácticas del lenguaje normativo o estándar con el fin de evitar

desviaciones y la omisión de información, propia de los sublenguajes y que puede causar problemas en la comprensión y en el procesamiento del lenguaje.

El esquema basado en Schwitter (1998 como se citó en Ramírez Polo, 2012), representa de forma gráfica los lenguajes controlados en el marco de la teoría general del lenguaje. Se observa cómo el lenguaje estándar o normativo es un subsistema del lenguaje natural. Asimismo, los sublenguajes se componen de un subconjunto de vocabulario y estructuras del lenguaje normativo o estándar, aunque también incluyen construcciones y vocabulario propios del lenguaje natural que se desvían de la norma. (p. 194)

Los lenguajes controlados están formados por elementos de los sublenguajes, ya que estos restringen su gramática y vocabulario de manera natural, pero al mismo tiempo contienen reglas impuestas para corregir las desviaciones y adaptarlas al lenguaje normativo o estándar.

En caso de esta investigación se utiliza la definición dada por Ramírez Polo:

Se define como lenguajes controlados al conjunto de elementos del lenguaje utilizado con fines o en dominios específicos, dirigido a un grupo de destinatarios concreto, que se caracteriza por presentar restricciones en el vocabulario, la gramática y el estilo. Tales lenguajes se utilizan fundamentalmente en la creación de documentación técnica debido, por un lado, a las características de este tipo de documentos y, por otro, al contexto en el que se producen. (Ramírez Polo, 2012)

Con respecto a las áreas de control en un lenguaje controlado, se distingue básicamente entre el control léxico y el gramatical. El control léxico o de vocabulario es el más común y es un elemento fundamental a la hora de restringir la redacción de los textos, de forma que solo se emplee vocabulario validado y definido previamente. Este control comprende dos niveles: por un lado, el control del vocabulario general o básico, destinado a evitar fenómenos como la sinonimia o la homonimia, y, por otro, el control de la terminología específica de algún área de conocimiento. En cuanto al control gramatical, abarca más categorías, como la sintaxis (estructura de las frases), la morfología, la ortografía y el estilo (pragmática).

¿Qué es un vocabulario controlado?

Un vocabulario controlado es un conjunto limitado de términos y reglas que forman parte de una estructura con fines de indexación y recuperación de información (COAR, «FAQsfor Controlled Vocabularies») que puede incluir términos preferidos y variantes (Bernal, «Uso de Vocabularios Controlados en Repositorios. La experiencia de DIGITAL.CSIC»). (Albuquerque, 2018)

Además se puede decir que los vocabularios controlados “son listas de términos válidos para denominar los conceptos de una disciplina, base documental, biblioteca, un repositorio, que se utilizan para etiquetar o indizar temáticamente” (Osés, 2017).

Sin embargo, normalmente el vocabulario controlado es más que una mera lista. Generalmente llevará incorporada alguna forma de estructura semántica, que estará diseñada para:

1. Controlar los sinónimos, eligiendo una de las formas como la aceptada y reenviando a ella desde todas las demás.
2. Distinguir entre homógrafos.
3. Juntar o enlazar aquellos términos cuyos significados muestren una relación más estrecha.

En términos generales, un vocabulario controlado puede oscilar entre una breve lista de términos claramente definidos, mutuamente excluyentes y exhaustivos que son las únicas opciones para el uso en un contexto específico o algo tan complejo como un tesoro con miles de términos y relaciones entre ellos.

Esta misma autora refiere que son beneficiosos en el proceso de indexación ya que los catalogadores utilizan el mismo término para referirse a un mismo concepto (por ejemplo, una persona, un lugar o una cosa) de una manera consistente. En la fase de recuperación, los vocabularios controlados guían a los usuarios en sus búsquedas si no conocen el término correcto para un concepto dado.

Según Barité (2014) para los defensores del control de vocabulario, la normalización opera favorablemente tanto sobre la indización de documentos (aportando una estructura de referencia para el indizador), como sobre la recuperación de información. Subyace la idea de que tales operaciones jerarquizan y mejoran la calidad de los términos de indización y las búsquedas temáticas, en contraposición al uso del lenguaje natural, no controlado ni estructurado; idea interpelada

cotidianamente por la mejora continua de las aplicaciones informáticas basadas en el lenguaje. (p. 101)

Es así como pueden distinguirse diferentes tipos de vocabularios controlados entre ellos los sistemas de clasificación bibliográficos, las listas de encabezamientos de materia las listas de autoridades, los mapas conceptuales, las taxonomías, las folksonomías, los anillos de sinónimos, las ontologías, y los tesauros entre otros, siendo los principales los dos mencionados primeramente y el último. En todos ellos se intenta presentar los términos tanto en orden alfabético como de forma sistemática.

I.3.3. Tipos de vocabularios controlados (Albuquerque, 2018):

➤ Encabezamientos de materia

Los encabezamientos de materia son listas alfabéticas formadas por encabezamientos y subencabezamientos que sirven para designar la temática a un conjunto de documentos. Los encabezamientos están formados por una o varias palabras que representan conceptos, tratando de condensar el tema sobre el que trata el documento.

Las funciones de las listas de encabezamientos son:

- Describir el contenido de los documentos indexados.
- Agrupar todos los documentos de temáticas afines.
- Evitar la ambigüedad expresando cada concepto en un término.
- Permitir la recuperación de los documentos por el campo materia.
- Relacionar documentos a partir de las relaciones semánticas entre términos.

➤ Catálogo de autoridades

Se define como autoridad a un conjunto de entidades normalizadas, compuestas por un registro y una clave. El registro de autoridad contiene información asociada con una entidad mientras que la clave de autoridad es un código que identifica unívocamente a la misma. Un catálogo de autoridades es un vocabulario controlado constituido por un conjunto de registros normalizados de autoridad. La práctica de mantener un catálogo de

autoridades, unificando los puntos de acceso a información normalizada se la llama control de autoridades. Este proceso tiene como objetivo:

1. Probar que dos valores son idénticos comparando por la clave de las autoridades.
2. Ayudar a completar metadatos con valores correctos.
3. Mejorar la calidad de los metadatos.
4. Mejorar la interoperabilidad compartiendo un nombre de autoridades con otra aplicación.
5. Reducir los tiempos de carga de metadatos.
6. Organizar la información.

Un registro de autoridad puede contar con manejo de variantes, es decir variaciones sintácticas de un registro dadas por delecteos y faltas de ortografía, mayúsculas frente a variantes en minúsculas, fechas diferentes.

➤ **Taxonomías**

Una taxonomía es un sistema de clasificación que permite agrupar un conjunto de elementos dentro de categorías predefinidas, llamadas taxones. Los taxones deben poder ser identificados unívocamente dentro de la taxonomía, pueden estar contenidos entre sí, o relacionadas de cualquier otra manera. Un ejemplo de taxonomía es la taxonomía biológica, utilizada para clasificar los seres vivos en reinos, especies y razas. En los sistemas de gestión de contenido, como Drupal y Wordpress, las taxonomías tienen un rol fundamental ya que permiten organizar el contenido en categorías y subcategorías, en el caso de Wordpress. En Drupal las taxonomías se utilizan no solo para organizar contenido, sino como base arquitectural para el desarrollo de otros módulos, como pueden ser menús.

➤ **Ontologías**

Una ontología es una definición formal de tipos, propiedades, y relaciones entre entidades cuyo fin es representar el conocimiento (entidades, ideas, eventos) limitando la complejidad del mismo a dominio en particular. Es considerado un vocabulario controlado ya que define los términos y las

relaciones para la comprensión de un área del conocimiento, así como las reglas para poder combinar los términos para definir las extensiones.

Las ontologías constan de términos (o clases), relaciones, propiedades, instancias y axiomas. Los términos son un conjunto de objetos (físicos, tareas, funciones, etc.) y las interacciones entre estos son representadas por las relaciones. Las relaciones más utilizadas son:

- ❖ Hiponimia: palabra cuyo significado está incluido en el de otra. Relación es-un (is-a). Por ejemplo: la palabra 'escritorio' es hipónimo de 'mueble' (siendo mueble hiperónimo de 'escritorio'), ya que posee todos los rasgos semánticos y añaden otras características para diferenciarlas de esta.
- ❖ Meronimia: una palabra que nombra una parte de un todo, relación part-of (parte-de). Por ejemplo: la palabra 'dedos' es un merónimo de 'manos' siendo 'manos' holónimo de 'dedos'.
- ❖ Sinonimia: relación que asocia dos términos que palabras que tienen un significado similar o idéntico entre sí. Por ejemplo, vehículo es un sinónimo de automóvil.

Los objetos que conforman una ontología se describen por medio de un conjunto de características o atributos llamados propiedades. Las instancias son la representación de los objetos. Por último los axiomas son teoremas que permiten definir aseveraciones que se cumplen siempre.

Existen tres tipos de axiomas: relacionales, no-relacionales y generales.

1.3.4. El control del vocabulario en el entorno digital.

Según Barité (2014) el control del vocabulario en ambientes digitales aún es motivo de interrogantes y discusión; de hecho, se escribe mucho sobre esta cuestión. En este marco, dos fenómenos concurrentes inciden en la evolución del debate entre lenguaje natural y control de vocabulario.

El primero de ellos es el desarrollo de nuevos sistemas de organización del conocimiento (manifiestamente las ontologías y las folksonomías) cuya estructuración interna pone en cuestión el modelo tradicional de creación de vocabularios controlados.

El segundo fenómeno que se ha generalizado en los últimos años es el desarrollo de herramientas y aplicaciones que operan como formas nuevas de 'control' lingüístico. Además de los menús que se abren a medida que se escriben letras en las ventanas de búsqueda, aparecen sugerencias cuando hay faltas de ortografía del tipo "quizás quiso decir", formas de búsqueda avanzada que mejoran la recuperación de datos, y el uso de comillas para refinar la búsqueda mediante expresiones compuestas.

Aún en este renovado y dinámico escenario, las operaciones de control de vocabulario y los vocabularios controlados tienen un importante rol a jugar, como herederos de una tradición normalizadora, homogeneizadores de lo heterogéneo, ordenadores de la producción documental en cualquier formato, compiladores y sistematizadores de la terminología de los campos especializados, jerarquizadores de conceptos y conocimientos.

Los vocabularios controlados se han nutrido del lenguaje natural. El nuevo desafío se sitúa ahora en la articulación de los probados y aprobados procesos de control de vocabulario con las nuevas aplicaciones e interfaces, y con las facilidades que ofrece la tecnología en el acceso y la organización de la información para lograr una contribución efectiva en lo que siempre ha sido su más reconocida fortaleza: la presentación racional y organizada de artefactos, conceptos y conocimientos fiables, actuales y pertinentes a los intereses de cada comunidad de usuarios.

1.3.5. Los vocabularios controlados y su uso en los repositorios.

El problema de la representación de los recursos en los repositorios se ha convertido en complejo poco a poco, principalmente por la diversidad de plataformas de software y esquemas de metadatos existentes para representar los recursos de diferentes tipologías en los repositorios. Por tanto, los esquemas de metadatos a través de su conjunto de elementos diseñados en una estructura formal y relacionada con la semántica, la sintaxis y la obligatoriedad en sus valores, pasan a ser un elemento importante en la representación de recursos dentro de un repositorio. (Texier, De Giusti, Oviedo, Villarreal y Lira, 2013)

Según Giusti, et al. (2013), algunos de los problemas dados en la representación de los recursos en los repositorios están basados entre otros aspectos en los vocabularios controlados. Uno de los más evidentes es que: muchos campos

presentes en los esquemas de metadatos necesitan ser estandarizados y puedan ser identificados, recuperados y usados de forma clara y precisa. Por ejemplo, tesauros (listas de palabras o términos empleados para representar conceptos) para definir las temáticas de los recursos (Eurovoc o DeCS), sistemas de clasificación conocimiento como el CDU (Clasificación Decimal Universal), idiomas, referencias geográficas, tipos de recursos, materias como los LCHS (Library of Congress Subject Heading), frecuencia de entrega (mensual, bimensual, trimestral), entre otros.

Otro es el de la redundancia.

Para evitar la redundancia y garantizar la integridad de la información en los vocabularios, se pueden considerar tres puntos de vista: representación, referencia y presentación.

- Forma de representación. Puede estar en una lista simple de elementos o en elementos relacionados, que se encuentran en una tabla de una base de datos, en un archivo XML con un esquema particular, en un archivo de texto, etc. Forma de referenciar. Se necesita una relación para distinguir de manera unívoca un elemento en un vocabulario determinado, por ello, se debe tomar la decisión entre: un metadato vacío con un valor adicional para la referencia, un metadato con valor del vocabulario replicado junto con un dato adicional para la referencia y un metadato con la referencia como valor.
- Forma de presentación. Es como el usuario estará observando y utilizando los metadatos en el portal, es decir, en formularios de carga, en la página de presentación de metadatos, en la exportación de recursos, etc. La forma puede ser simple, intuitiva, e internacionalizable.

Es así como la utilización de los vocabularios controlados para describir los metadatos bibliográficos en los repositorios asegura que todos utilicen el mismo término para significar lo mismo, mejoran la interoperatividad entre repositorios y entre otros sistemas. Los vocabularios capturan la riqueza de las variantes terminológicas y promueven la consistencia a través de la elección de términos preferidos y la asignación de un mismo término para significar un contenido similar.

I.4. Herramientas de indización

Desde hace mucho tiempo para la realización del proceso de indización se han utilizado diversas herramientas que van a estar relacionadas con el tipo de indización que se realice en la unidad de información, entre los que podemos citar:

- Tesauros: es una lista estructurada de conceptos, destinados a expresar de manera unívoca el contenido de los documentos y de las consultas dentro de un sistema documental determinado, y a ayudar al usuario en la indización de los documentos y de las consultas.
- Índices: Signo autorizado que representa un concepto determinado y se utiliza para registrar los resultados del procesamiento analítico de la información.
- Listas de encabezamientos de materias: son lenguajes alfabéticos de materia y caen dentro de los lenguajes precoordinados enumerativos, que ordenan alfabéticamente los términos y no ofrecen posibilidades para coordinar clases más complejas en el momento de la indización.
- Ontologías: plantea que una ontología es una representación formal y explícita de la estructura conceptual de un campo del conocimiento. Una ontología es el soporte semántico de las palabras que son descritas como objetos lingüísticos en una base de datos léxica o terminológica. Las relaciones conceptuales que son representadas en una ontología son muy variadas y dependientes del campo del conocimiento a estructurar. Una ontología se construye a fin de compartir y reutilizar la información almacenada, la cual, por estar formalizada puede ser interpretada tanto por personas como por programas informáticos.

I.5. Los Tesauros

Origen del tesaurus

Para Soler Monreal (2009), la palabra *thésauroses* de origen griego, de la que derivó el término latino *thesaurus*, que significa tesoro, almacén de algo valioso.

El término **tesaurus** se utilizó por primera vez en la década de 1950, cuando tanto Howerton como Helen Brown utilizaron la palabra *tesauros*, aunque algunas publicaciones atribuyen este primer uso a Joyce y Neddham en 1956 y otros a Farradane en 1952 (Currás, 1991).

El sistema *Uniterm*, creado por Mortimer Taube en 1951 se considera muy relacionado con la aparición del tesauro. Este sistema representaba las materias con palabras llamadas *uniterms*, extraídas de los documentos, pero sin ningún tipo de control. A las palabras clave principales Taube les dio el nombre de “descriptores”, lo que según Currás (1991), supuso la acuñación del término. Con los descriptores se construían índices; por consiguiente, se llamó *indización* al proceso de obtener aquellas palabras de los documentos con las cuales se confeccionaban los índices. No se tiene noticia de la primera utilización del término **indización**, pero Currás (1991) lo sitúa hacia 1945.

Con anterioridad al sistema *uniterm*, en 1947, Calvin N. Moers presentó el sistema *Zator*, basado en descriptores, simples o compuestos, que posibilitaban la combinación de conceptos.

El primer tesauro con la función de control del vocabulario fue el elaborado por el *Engineering Information Center* de *E.I. Dupont de Nemoursen* 1959.

En 1960, el principal Centro de Información del Departamento de Defensa de Estados Unidos redactó su primer tesauro. Un año después, el *American Institute of Chemical Engineers* publicó el *Chemical Engineering Thesaurus*, basado en el trabajo de Dupont. A partir de él se elaboró el *Thesaurus of Engineering Terms*, publicado por el *Engineering Joint Council* en 1964, que cubría toda la ingeniería. Los tesauros que se construyeron posteriormente estuvieron relacionados con la ciencia y la ingeniería. En 1967, se publicó en Gran Bretaña el *TEST Thesaurus of Engineering and Scientific Terms*, en el que después se basaría la norma británica 5723, que la ISO tomó posteriormente como modelo para la elaboración de la norma 2788 en su primera edición, en 1974. En 1970, se constituyó en la República Federal de Alemania un grupo de trabajo para construir un tesauro para la industria química. (p.18)

Existen diversas definiciones del término proporcionadas por renombrados documentalistas en 1971 (como se citó en Soler Monreal, 2010, p. 13-16), no obstante la investigación se basó en la siguiente:

Según Soler Monreal (2009) los tesauros son vocabularios controlados, de estructura combinatoria, definidos a priori, es decir, fijados con anterioridad, compuestos por

términos que reflejan conceptos que se relacionan entre sí semántica, jerárquica y asociativamente, que tienen como finalidad el control de sinónimos, y se utilizan para describir de manera unívoca, traduciendo del lenguaje natural al documental, el contenido de los documentos, para su posterior recuperación en un sistema documental dado, con el fin de satisfacer las necesidades de información. (p. 17)

I.5.1. Funciones del tesoro

Gil Urdicián (2004 como se citó en Soler Monreal, 2009) resalta tres funciones principales de los tesauros:

1. Normalizar el vocabulario. Es decir, controlar los accidentes que puede generar el vocabulario, que son la sinonimia y la polisemia, así como el género y el número de las expresiones.
2. Inducir al usuario hacia posibles alternativas, términos en los que el usuario no había pensado, “gracias a una serie de referencias cruzadas que indican las relaciones asociativas, jerárquicas o preferenciales que se pueden dar entre la terminología que lo compone”.
3. Representar los conceptos presentes en los documentos. Esta función es compartida con el resto de lenguajes documentales. (p. 17)

I.5.2. Objetivos del tesoro.

Según Lancaster (2002 como se citó en Soler Monreal, 2009) los objetivos de los tesauros pueden resumirse en dos:

- ❖ “Facilitar la representación consistente de las materias por parte de indizadores y usuarios que recuperan, evitando la dispersión de los elementos relacionados. Esto se consigue con el control (agrupación) de los sinónimos y cuasisinónimos y la distinción de los homógrafos”.
- ❖ “Facilitar la realización de una búsqueda amplia sobre una materia enlazando los términos con relaciones paradigmáticas o sintagmáticas”. (p. 18)

El objetivo más importante sería conseguir la consistencia en la indización:

“será más probable que los indizadores estén de acuerdo sobre los términos necesarios para representar determinada materia, si éstos son seleccionados de una lista previa, que si pueden utilizar cualquier término. Lo mismo ocurre en el proceso de búsqueda: será más fácil identificar los términos apropiados

a una necesidad de información si son seleccionados de una lista determinada. Por lo tanto, el vocabulario controlado facilita la coincidencia entre el lenguaje de los indizadores y el de los usuarios que hacen la búsqueda”. (Lancaster, como se citó en Soler Monreal, 2009, p.18)

I.5.3. Normas y directrices para la elaboración de tesauros.

Para Soler Monreal (2010, p. 20-23) el primer especialista que propuso unas normas para la construcción de encabezamientos alfabéticos de materias fue el norteamericano Charles Ammi Cutter en 1876: *Rules for a Dictionary Catalogue*.

Las normas nacionales americanas para la construcción de tesauros, la ANSI Z39.19 *Guidelines for the construction, format and management of monolingual controlled vocabularies*, publicadas por el *American National Standards Institute* aparecieron por primera vez en 1974. Estas normas fueron revisadas a finales de los años 70, y en 1980 apareció la segunda edición. Con la posterior edición de 1993 se amplió de modo significativo el ámbito de aplicación de la norma con la indicación *format and management*. Ésta ofrece más detalles respecto al control terminológico y las presentaciones en pantalla. Esta tercera edición es análoga al estándar internacional ISO 2788 y la norma británica BS 5723. Una nueva revisión fue aprobada en 1998, a pesar de que fue muy criticada por no recoger la evolución de los sistemas informáticos (Franchini, 2005).

En julio de 2005 se publicó la cuarta edición, la norma ANSI/NISO Z39.19:2005, surgida del *Workshop on Electronic Thesauri*, promovido por NISO y celebrado los días 4 y 5 de noviembre de 1999, con la finalidad de redactar una norma que abarcara la elaboración de tesauros electrónicos.

Por su parte, la *British Standard Institution*, BSI, elaboró una norma británica, la BS 5723, publicada por primera vez en 1979, equivalente a la norma ISO 2788. La última revisión se ha publicado en cinco partes: las dos primeras vieron la luz en el año 2005, las dos siguientes en 2007 y la última parte está en preparación. Se trata de la norma BS 8723 *Structured vocabularies for information retrieval*.

La primera normativa internacional, la ISO 2788, de la *International Organization for Standardization*, apareció en 1974, basada en las recomendaciones de la Unesco. A partir de la norma ISO, y equivalente a ella, se desarrollaron normas en distintos

países a través de sus organismos de normalización como en Francia AFNOR (*Association Française de Normalisation*), en Alemania DIN (*Deutsches Institut für Normung*), en Italia UNI (*Ente Nazionale Italiano di Unificazione*), en Cuba ONN (Oficina Nacional de Normalización) y en España UNE.

Por otro lado, las recomendaciones de la Unesco para la elaboración de tesauros multilingües publicadas en 1976 y 22 revisadas en 1979, dieron lugar a la norma ISO 5964, aparecida en 1985. Por su parte, la IFLA publicó en 2005 *Guidelines for multilingual thesauri*, incorporando los cambios propuestos por la ANSI/NISO Z39.19:2005.

La revisión de la norma británica, el estándar BS 8723, ha sido propuesta como soporte del trabajo de normalización internacional. En abril de 2007 se presentó el proyecto ISO NP 25964, que propone la revisión de las normas ISO 2788 e ISO 5964 (sobre tesauros monolingües y multilingües respectivamente). En el proyecto participan nueve países: Francia, Alemania, Canadá, Finlandia, Nueva Zelanda, Suecia, Reino Unido, Ucrania y Estados Unidos. Su publicación está prevista para 2010. Esta norma se publicará en dos partes:

Parte 1: *Thesauri for information retrieval*, que recogerá todas las directrices sobre tesauros, ya sean mono o multilingües. (2011)

Parte 2: *Interoperability witho ther vocabularies*, que incluirá los vocabularios distintos de los tesauros y su interoperabilidad. (2013)

1.5.6. Existen algunas condiciones que por lo general debe cumplir todo tesoro:

- ❖ Debe ser un lenguaje especializado
- ❖ Debe estar normalizado
- ❖ Las unidades lingüísticas que lo componen, por tratarse de léxicos dedicados a un tema concreto, adquieren la categoría de términos; convertidos en palabras claves, por lo que determinan el tema del que trata un documento.
- ❖ Esta palabra clave se relaciona entre sí, jerárquicamente, de forma asociativa por semejanzas de equivalencia.
- ❖ Estos procesos de relación se podrán realizar de forma precoordinada o postcoordinada.

- ❖ Se trata de lenguajes terminológicos, usados con fines documentarios, por lo que se convierten en leguajes documentarios utilizados en los procesos de indización o clasificación y en la recuperación de la información.
- ❖ Deben permitir la introducción o eliminación de términos para mantener su actualidad constantemente.
- ❖ Deben servir para convertir el lenguaje natural, de los documentos, en lenguaje concreto, normalizado, que permita controlar la información contenida en el documento.
- ❖ Ha de servir como puente entre el documento y el usuario, donde el especialista en información es el eslabón principal.

Aunque cumplen estas características comunes, también tienen diferencias. Por lo que existen distintas clases de tesauros:

- Según su ámbito temático:
 - Macrotesauros: son aquellos donde están implicadas varias áreas temáticas.
 - Microtesauros: son las partes o áreas temáticas implicadas en un macrotesauro.
- Según el idioma:
 - Multilingües: se reconocen varios idiomas en las correspondencias.
 - Monolingües: un solo idioma.
- Según la ordenación:
 - Alfabéticos.
 - Sistemáticos: por materias y facetas. (Soler Monreal, 2009, p. 20-23)

I.5.7. Estructura de los tesauros.

Los tesauros están constituidos por unidades léxicas.

Unidades léxicas

Descriptores y no descriptores y relaciones semánticas: relaciones de equivalencia, relaciones jerárquicas y relaciones asociativas; que pueden comportar las siguientes categorías: grupos de descriptores que agrupan los términos de indización bien por campos (temas), bien por clases de términos (facetas).

Descriptores

“Término (palabra o expresión), que se ha escogido, a partir de un conjunto de sinónimos, de cuasisinónimos y de términos emparentados, para representar de manera unívoca, un concepto susceptible de intervenir en los documentos y en las consultas que se examinan dentro de un sistema documental dado”.

“Palabra o grupo de palabras incluidas en un tesoro y escogidas de entre un conjunto de términos equivalentes para representar sin ambigüedad una noción contenida en un documento o en una petición de búsqueda documental”.

Existen varias clases de descriptores:

Descriptores simples

“Son aquellos que están formados por un solo término y a menos que no tengan un significado unívoco estarán en la misma situación operativa de los unitérminos.”

Descriptores compuestos

“Son aquellos que para representar un concepto requieren de varias palabras. También algunos los llaman sintagmáticos. En los últimos tiempos los descriptores simples se han visto desplazados por los descriptores sintagmáticos debido a que estos permiten profundizar en los conceptos y satisfacer, de este modo, las necesidades más específicas dadas por la extensión bibliográfica en los temas y también eliminan la ambigüedad que conllevan los unitérminos de Taube.”

Existen 2 métodos de formulación de descriptores compuestos o sintagmáticos:

1. Unión morfológica (o sintáctica para otros autores) en que 2 o más términos se unen utilizando preposiciones o artículos o, por la adjetivación de uno de ellos.
2. Unión lexicológica (o semántica), según la cual 2 o más términos se funden en uno.

El lenguaje descriptor involucra listas de términos mucho más amplias que las requeridas por el lenguaje de unitérminos o palabras clave y muestra las ventajas que hacen su uso de más aplicación en el campo del tratamiento documentario, especialmente si se encuentra estructurado.

Todos los términos descriptores se relacionan semánticamente en tres dimensiones:

- la sinonimia
- la polisemia

- la antonimia

La sinonimia o relación de equivalencia entre el término descriptor y el sinónimo no descriptor. Se produce cuando un solo concepto está representado por varios significantes. Se soluciona con la existencia convencional de descriptores y no descriptores aprobados. El vocablo más usado será el término descriptor y los demás los no descriptores.

Polisemia: se trata de la existencia de varios significados para un solo significante. Es uno de los problemas más complejos en la definición de términos. Se distinguen distintas formas de polisemia:

- Polisemia pura: hay significantes iguales que conservan relación semántica pero se refieren a significados distintos. Tal podría ser el caso del término suspenso, aplicado a los evaluados en un examen y a los archivos.
- Homonimia: hay significantes iguales que hacen referencia a conceptos totalmente distintos sin relación.
- Homofonía: los significantes son distintos en grafía, sin embargo su pronunciación es igual refiriéndose a significados distintos.

Antonimia: es el empleo controlado de términos opuestos.

- No descriptores: "Término (palabra o expresión), incluido dentro de un tesoro, tomado de una lista de sinónimos, de cuasisinónimos y de términos emparentados con uno o dos descriptores de ese tesoro, que está ligado a tal(es) descriptor(es) por una relación de equivalencia semántica susceptible de intervenir en los documentos o en las consultas, pero no utilizable para indizar esos documentos o formular esas consultas".

"Palabra o conjunto de palabras incluidas en un tesoro con prohibición de uso y reenvío a uno o más descriptores utilizables".

Relaciones semánticas: son todas aquellas relaciones que se establecen entre los términos de un tesoro. Podemos distinguir las siguientes:

- Relaciones de equivalencia: se derivan de la gran proximidad fonética entre los descriptores y no descriptores (sinónimos, cuasi-sinónimos, etc). Este tipo de relaciones se suelen representar mediante el término USE (relación entre el

término no preferente y el término preferente, o término a usar más aceptado) y las siglas UP (Usado por).

- Relaciones jerárquicas: son relaciones asimétricas, establecidas entre los descriptores, en la que un término es superior o genérico de otro. Esta estructura clasificatoria distingue al tesoro de las listas alfabéticas de materias o palabras clave, así como de diccionarios y léxicos. De esta forma se crean campos conceptuales que clasifican los términos y los agrupan en un campo delimitado de significados o facetas que se van agrupando en disciplinas y sub-disciplinas.

Para la presentación de las relaciones jerárquicas se suelen utilizar determinadas siglas como T.G. (Término Genérico) y T.E. (Término Específico); si un descriptor tiene un sólo descriptor por encima de él es monojerárquico, si tiene dos o más es polijerárquico.

Las relaciones jerárquicas pueden ser de dos tipos:

Relaciones género-especie: se aplica a acciones, propiedades y agentes.

Relaciones parte-todo: el nombre de una parte implica y evoca la del todo poseedor. Se aplica a objetos, lugares, disciplinas y sus partes de estudio, estructuras sociales.

- Relaciones asociativas: son las relaciones asimétricas establecidas entre descriptores de distintas familias, distintas a las jerárquicas o de equivalencia, y que designan conceptos capaces de evocar mutuamente, por asociación de ideas, relaciones entre ellos. La mayoría de las relaciones asociativas se establecen entre términos que comparten un lexema común o que mantienen unas relaciones de hiperonimia o hiponimia.

También podemos añadir Notas de Alcance (N.A.) con el fin de ofrecer una explicación o descripción del significado del término.

I.5.8. Métodos para la construcción de tesauros.

Existen varios métodos para la creación de los tesauros descritos en (López Alonso, s.f. como se citó en Hernández Quintana, 2004). Pero por las características de esta investigación se utilizó el siguiente:

- I. Elección del área temática y razones para su elaboración.
- II. Comprobación de la no existencia de otro tesoro similar, tanto en ámbito nacional como europeo.

Van Slype (s.f. como se citó en Hernández Quintana, 2004) describe el proceso de adaptación de un tesoro de la forma siguiente:

1. Se adquiere los tesauros disponibles sobre la temática.
2. Se prepara una muestra de documentos para indizar,
3. Se indizan los documentos en lenguaje natural,
4. Se intentan traducir los descriptores libres de la indización en descriptores controlados, extraídos de cada uno de los tesauros examinados, uno tras otro.

El tesoro que haya proporcionado la mayor cantidad de descriptores, que respondan a las necesidades del sistema documental, será el candidato para la elección.

Por debajo del 95 y más del 80 por ciento de los descriptores necesarios, deberá ser complementado por los descriptores que falten para llegar a ser el adecuado.

Por debajo del 80 por ciento deberá construirse un nuevo tesoro para el sistema documental en estudio, utilizando los descriptores como fuentes de terminología del tesoro a preparar.

- III. Delimitación del tema principal, sus auxiliares y marginales.
- IV. Recopilación de los términos de:
 - Vocabularios, diccionarios, libros, y otras fuentes secundarias (“a priori”).
 - Publicaciones periódicas y otras fuentes primarias (“a posteriori”).
- V. Normalización y depuración de vocabulario.
- VI. Agrupación de los términos y estructuración de las relaciones jerárquicas, asociativas y de equivalencia.
- VII. Puntualización de las notas de aplicación.
- VIII. Edición

El tesoro debe ir precedido de una información detallada que indique claramente su objetivo, estructura y los campos que cubre, teniendo en cuenta las consideraciones internas referidas al propio centro para el que se desarrolla. También, es necesario indicar brevemente las reglas que

han surgido su elaboración, así como las fuentes y métodos que se han usado en la selección de los descriptores, y las fechas previstas para su actualización.

I.5.8.1. Herramientas de gestión de Tesauros

Existen varias herramientas informáticas para la edición de tesauros y el control de vocabularios descritas en la página <https://datos.gob.es/es/noticia/herramientas-para-la-gestion-de-vocabularios-controlados-con-soporte-skos> .

Entre las mencionadas se escogió TemaTres.

TemaTres (de Albuquerque, 2018):

TemaTres es una herramienta Web que permite la creación y gestión de vocabularios controlados, tesauros, taxonomías y otros modelos de representación formal del conocimiento. Permite representar, publicar y exponer vocabularios controlados en varios esquemas de metadatos como pueden ser SKOS-Core, Dublin Core, MADS, entre otros.

Esta herramienta dispone de un workflow de gestión de términos que establece qué puede hacerse con un término según sea su estado, y cuáles son las condiciones que debe reunir para poder pasar al siguiente. Los estados por los que pueden atravesar dichos términos son Candidato, Aceptado y Rechazado.

Un término puede ser eliminado en cualquiera de estos tres estados, salvo que el mismo se encuentre en aceptado y no sea término libre. Por otra parte, sólo los términos aceptados pueden tener relaciones con otros términos. También es posible representar las relaciones jerárquicas entre los términos, indicando si estos son Términos Genéricos (TG), Términos Específicos (TE), Términos No-preferido (UP) y Término Relacionado (TR).

La integración de vocabularios TemaTres dispone de dos modelos posibles para la implementación de vocabulario multilingüe:

- Modelo centrado con un vocabulario de referencia local
- Modelo federado con vocabulario de referencia externo
- Modelo centrado con un vocabulario de referencia local

En este modelo TemaTres contiene y gerencia en el mismo vocabulario los términos propios del vocabulario fuente y las relaciones de equivalencia hacia los términos de

los vocabularios de destino. Esta opción puede resultar útil si el vocabulario de destino no existe o sus servicios terminológicos no están disponibles, o las correspondencias terminológicas resultan incompletas.

Se puede concluir entonces que:

- Por lo general, los tesauros se presentan con una estructura tripartita, esto es, un apartado jerárquico en donde se presentan los descriptores en el contexto de su campo semántico, una parte alfabética para facilitar la entrada en el tesoro por la primera palabra significativa de descriptores y no descriptores y que expresa las relaciones del vocabulario, y un índice permutado que permite el acceso por el resto de términos significativos.
- Los descriptores constituyen las unidades básicas que conforman los tesauros. Su función principal es describir un concepto en forma unívoca en el contexto de un sistema de información. El conjunto de unidades lingüísticas denominadas descriptores y sus relaciones correspondientes conforman el tesoro, herramienta que se utiliza para indizar de manera normalizada la información que contienen los documentos.

CAPÍTULO II: METODOLOGÍA UTILIZADA PARA LA ELABORACIÓN DEL TESAURO ESPECIALIZADO EN EL ÁREA DE LA CIENCIA DE LOS MATERIALES.

II.1. Caracterización general del Repositorio Nínive.

El Repositorio de la Universidad de Moa Dr. Antonio Núñez Jiménez, está basado en la plataforma de software libre DSpace, y es un servicio digital para favorecer la recogida, conservación y distribución de los recursos digitales relacionados con el aprendizaje, la enseñanza y la investigación; elaborados por sus miembros, bajo una política de Acceso Abierto. Debe su nombre a la "primera biblioteca sistemáticamente recopilada" de la que se tiene noticias como conjunto organizado de libros y documentos, del rey asirio Asurbanípal, descubierta en el siglo XIX al excavar en Nínive.

El repositorio está dirigido a la comunidad académica de la institución y de otras interesadas. El proyecto de repositorio responde a sectores muy específicos como la educación superior o universitaria y otras áreas de la investigación científica. Se encuentra organizado siguiendo la estructura académica y científica de la universidad y del Ministerio de Educación Superior (MES), por facultades.

La "Facultad de Metalurgia y Electromecánica" está compuesta por subcomunidades donde se puede encontrar el Centro de Estudios de Energía y Tecnología de Avanzada de Moa, Departamento de Eléctrica, Departamento de Mecánica y el Departamento de Metalurgia-Química.

II.2 Alcance temático del tesauro.

Para el desarrollo del trabajo se contempló el dominio temático aplicable en la Ciencia de los Materiales, presente en la Facultad de Metalurgia-Electromecánica. Es un campo multidisciplinario que maneja diferentes disciplinas, y los aplica en varias áreas de la ciencia y la ingeniería todas ellas de larga tradición como la: Metalurgia, Física y Química, entre otras, donde esta concurrencia ha originado tensiones entre conceptos y terminología.

La investigación es de tipo aplicada, debido a que se dirige hacia la construcción de un producto terminológico que brinda una solución para la indización y recuperación

de los documentos en el Repositorio Institucional Nínive de la Universidad de Moa Dr. Antonio Núñez Jiménez en el campo temático referido.

Se escoge esta área temática ya que no se dispone de una herramienta léxica adecuada al contexto de la institución que permita establecer la terminología que represente a la colección de esta área temática. Además es la colección más grande y compleja desde el punto de vista de la terminología que maneja, lo que dificulta el proceso de indización para las especialistas que trabajan en esta especialidad.

II.3 Población objeto de estudio.

La población objeto de estudio de esta investigación son todos los documentos de la colección de la Facultad de Metalurgia y Electromecánica y los usuarios que investigan, estudian, leen o trabajan en temas afines con la Ciencia de los Materiales que harán uso del Repositorio.

II.4. Guía metodológica de trabajo para la confección del tesoro.

Primera Fase

1.a) Se realizó el estudio teórico y metodológico acerca del tema en cuestión, así como se realizaron amplias búsquedas bibliográficas sobre la construcción de algunos tesauros, estructura, formato y software utilizados.

En este caso como guía y ejemplo para esta investigación se utilizaron los siguientes documentos:

Tabla 1. Documentos guía que sirvieron para la construcción del tesoro.

No.	Título del documento	Responsabilidad	Disponible en:
1.	Construcción de un tesoro especializado en Comunicación Social para el contexto colombiano.	Pinzón Otálora, Oscar J.	http://repository.lasalle.edu.co/

2.	Propuesta preliminar de diseño de un Tesouro de la Música Cubana para uso de instituciones de información especializadas en esta materia.	Rodríguez Noguera, Marbelys	http://BVIRTUAL/biblioteca/tesis/Tesis202006/tesau.pdf
3.	Propuesta para la construcción de un Tesouro Especializado en Arquitectura.	Malavassi Aguilar, Flora A.	http://repositorio.sibdi.ucr.ac.cr:8080/jspui/bitstream/123456789/253/1/28444%20Tom%20l.pdf

Al finalizar este paso quedó decidido:

- Tipo de método para aplicar en la construcción del tesouro.

La metodología utilizada fue con un procedimiento mixto ya que se emplearon aspectos del método de comité y el método empírico o como también se le conoce de abajo a arriba.

El método de comité fue aplicado a partir de la creación de un grupo de trabajo de especialistas en las distintas materias que alcanzó el tesouro.

El método empírico o de abajo a arriba fue aplicado utilizando la terminología de base generada en la propia práctica del proceso de indización de los documentos que forman la colección de la Facultad de Metalurgia y Electromecánica y teniendo en cuenta además el comportamiento de los usuarios de esta facultad al formular sus solicitudes de información en el Centro de Información de la institución.

Este método tiene autoridad bibliográfica, lo que significa que un término está autorizado si se utiliza en la literatura o al menos aparece con bastante frecuencia.

- El tipo de tesouro en cuanto a la materia tratada.

Es un tesouro especializado en el dominio temático de la Ciencia de los Materiales. Dividido a su vez en cuatro microtesauros de las áreas de la mecánica, metalurgia, eléctrica y tecnología ambiental.

- Tipo de tesouro por su cobertura.

Se presenta un tesoro de tipo institucional, ya que se construyó a partir de la indización realizada a la colección documental generada en la Universidad de Moa Dr. Antonio Núñez Jiménez para ser aplicado al Repositorio Institucional Nínive.

- Tipo de tesoro por su estructura.

Se considera un tesoro de tipo jerárquico, ya que parte de una esquematización que muestra las relaciones jerárquicas entre los términos, desde el término genérico hasta los específicos.

- Por su forma al presentar los términos.

El tesoro tiene presentación alfabética y sistemática.

- Proceder para recopilar las unidades léxicas.

Para recopilar las unidades léxicas se trabajó con un total de:

- 921 documentos pertenecientes a la Facultad de Metalurgia-Electromecánica.
Desglosados en: 897 tesis y 24 artículos.

De estos documentos se tomaron las palabras claves tal y como fueron asignadas durante el proceso de indización, las cuales después fueron objeto de intervención semántica y lingüística. Estas palabras claves se gestionaron a partir de un documento en Word para su posterior revisión y análisis por parte de los especialistas.

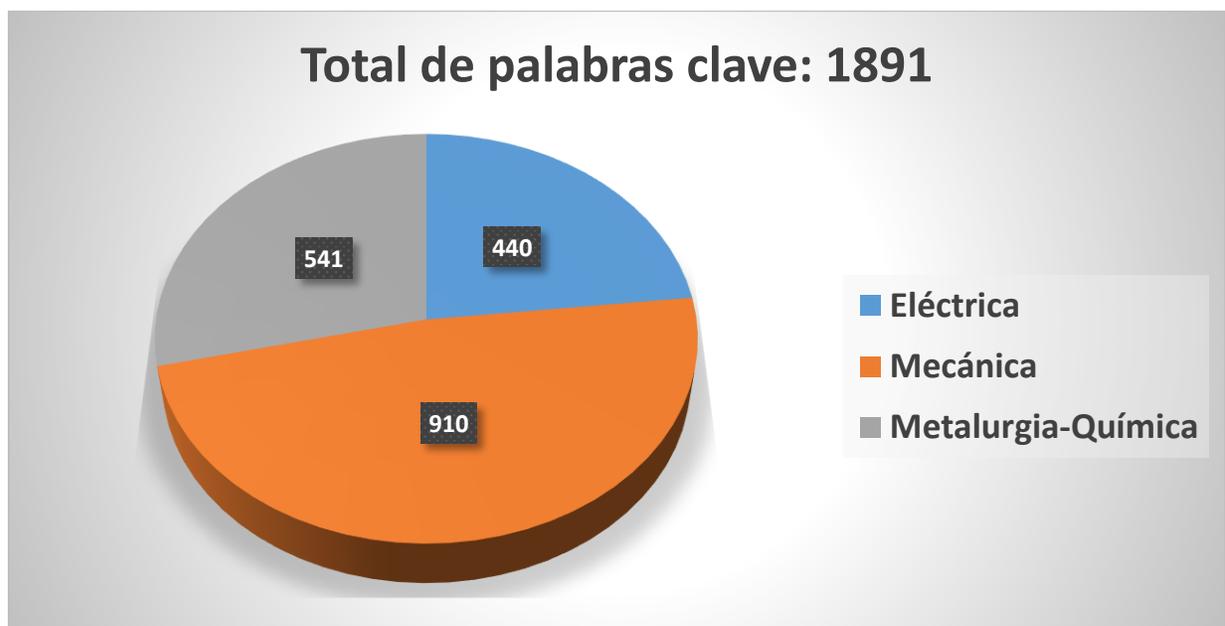


Fig.1. Segmentación de las palabras claves utilizadas para generar el tesoro.

- Gestión de los términos a partir de un software.

Para la construcción del tesoro se utilizó un software libre, en este caso fue el software TemaTres.

1. b) Determinar los subcampos de conocimientos que abarcó el tesoro a partir de la temática general.

Para determinar las principales temáticas del tesoro y los términos específicos de cada una de ellas se utilizó la Clasificación Decimal Dewey, en su parte de las ingenierías y temas afines basada en la 18 ed. con adiciones de la 19, del 1980 en su versión impresa y en su versión digital de la edición 20 de 1995. Esta herramienta es la utilizada en Cuba por el Sistema de Bibliotecas Universitarias y otros sistemas de información. Se utilizó el Rubricador del Sistema Nacional de Información Científica y Técnica (repertorio de la Academia de Ciencias de Cuba), el criterio de los especialistas del dominio temático y el criterio de las especialistas del Centro de Información que realizan el proceso de indización.

Los campos léxicos del tesoro quedaron establecidos de la siguiente manera:

- Ingeniería eléctrica
- Energética
- Electrotecnia y electrónica
- Luminotecnia

TG Ingeniería mecánica

- Ciencia de los materiales
- Máquinas, piezas y herramientas
- Mecánica de los gases
- Mecánica aplicada
- Mecánica automotriz
- Mecánica de fluidos
- Mecanizado de materiales
- Resistencia de materiales
- Sistemas de propulsión

- Termodinámica

TG Ingeniería metalúrgica y química

- Beneficio de las menas
- Electroquímica aplicada
- Elementos químicos
- Industria metalúrgica
- Materiales cementicios suplementarios
- Metalurgia, física y química
- Procesos químicos
- Propiedades mecánicas
- Química analítica
- Química industrial
- Reacciones y síntesis químicas
- Técnicas de análisis

TG Modelos científicos

- Modelación

TG Tecnología ambiental

- Aguas residuales
- Biodigestores
- Biogás
- Biomasa
- Reciclaje
- Ruido
- Seguridad industrial
- Tecnologías limpias
- Tratamiento de aguas residuales
- Tratamiento del agua

1. c) Determinar la cobertura conceptual.

Para la construcción del tesoro se realizaron búsquedas en el Repositorio Institucional Nínive por las diferentes comunidades que representa el tesoro y se fueron extrayendo todas las palabras claves conformadas en el índice de materia. Se tuvo en cuenta no sólo la existencia de las palabras claves utilizadas, sino la posibilidad además de que estas aparecieran como descriptor en algunas de las herramientas léxicas utilizadas, el cual se tomaría directamente de éste, si se considerara conveniente hacerlo.

Para la selección de las palabras claves se decidió tomar como criterios fundamentales:

- ✓ Escoger solo aquellas palabras que pertenecieran al campo de la Ciencia de los Materiales.
- ✓ Y las que fueran conceptos específicos aplicados a casos muy particulares de investigación, descartando aquellos términos que pertenecen a la terminología de otras disciplinas.

Es necesario destacar que en esta investigación se incluyeron algunos términos de materia llamados identificadores o descriptores auxiliares, por ser muy utilizados en las búsquedas de información solicitadas por los usuarios de este campo disciplinario en la institución. Se pueden citar en este caso nombres propios de métodos, transportes, equipos y técnicas.

Finalmente se conformó un documento Word en el cual se tabularon todas las palabras clave capturadas, con el fin de ser entregadas a los diferentes especialistas que se encargarían de su análisis y revisión.

Materia departamento de Eléctrica	Campo semántico
<p>Accionamiento electrohidráulico T.G. Accionamiento eléctrico TR Accionamiento de bombas</p> <p>Accionamiento eléctrico Accionamiento hidráulico Bombas Electroválvula</p>	Electrotécnia y Electrónica
<p>Motores asincrónico UP Máquinas asincrónicas TE Motor de inducción trifásico TR Bomba centrífuga Contactor magnético Control de velocidad Freno dinámico Generación de energía eléctrica Generador asincrónico Método del análisis espectral Método del momento en el entrehierro [MME] Motor de corriente alterna Sistemas de bombeo</p>	Electrotécnia y Electrónica

Fig. 2. Ilustración de la captura de las palabras claves en el documento de Word.

1.d) Herramientas bibliotecológicas utilizadas para la construcción del tesauo.

Para la construcción del tesauo se utilizaron varios documentos rectores de la bibliotecología:

Tabla 2. Herramientas bibliotecológicas utilizadas para la construcción del tesauo.

No.	Título del documento	Responsabilidad	Disponible en:
1.	NC 39-14 1983. Elaboración del Tesauo Unilingüe de Búsqueda Informativa.	Oficina Nacional de Normalización.	Biblioteca Universidad de Moa "Dr. Antonio Núñez Jiménez".
2.	Rubricador del Sistema Nacional de Información Científica y Técnica	Academia de Ciencias de Cuba	Biblioteca de la Universidad de Moa "Dr. Antonio Núñez Jiménez".
3	Sistema de Clasificación Decimal Dewey.		Biblioteca de la Universidad de Moa "Dr.

Antonio Núñez Jiménez".

1.e) Herramientas léxicas utilizadas para validar los términos.

Para la confección del tesoro se tuvo en cuenta obras de referencias especializadas tales como glosarios y diccionarios, entre otros recursos. Estos materiales ayudaron a precisar los términos y a esclarecer los significados de algunos de ellos para la redacción de las notas de alcance.

Tabla 3. Listado de los recursos consultados.

Diccionarios				
No.	Título	Responsabilidad de la obra	Año	Ubicación
1	Diccionario técnico de Mecánica de automóviles 1ra Edición	Colección Automotriz "Fenusa"		https://www.bcn.cl/obtienearchivo?id=documentos/10221.1/61186/1/207689.pdf&origen=BDigital
2	Diccionario Enciclopédico de la Siderurgia	Vázquez López, Jerónimo	1965	Biblioteca de la Universidad de Moa
3	The free dictionary	ByFarlex	2003	https://es.thefreedictionary.com/
4	MotorGiga	RedGiga	1998	https://diccionario.motorgiga.com/
5	Diccionario de Ingeniería Eléctrica	Universidad de Salamanca Escuela Técnica Superior de Ingeniería Industrial		http://electricidad.usal.es/Principal/Circuitos/Diccionario/Inicio

6	Diccionario de Ingeniería	de	Real Academia de Ingeniería		http://diccionario.raing.es/
Tesauros					
No.	Título	Responsable	Año	Ubicación	
1	Tesauro Eurovoc	Comunidades Europeas	2006	http://www.bizkaia.eus/descargar_documento.asp?url=kultura%2Feurovoc%2Fidb%2Fpdfs%2Fpermut_es.pdf&param=1	
2	Tesauro de Ingeniería Sanitaria y Ambiental 17 Ed.	Biblioteca Virtual de Desarrollo Sostenible y Salud Ambiental	2004	http://www.bvsde.paho.org/bvsair/e/manuales/esa/teses.pdf	
3	Tesauros del Patrimonio Cultural de España	Red Digital de Colecciones de Museos de España	2019	http://tesauros.mecd.es/tesauros/materias/1188762.html	
4	Tesauro de la UNESCO		1995	http://vocabulary.unesco.org/browser/thesaurus/es/	
5	Tesauro de Serbiula	Biblioteca Universidad de los Andes		http://www.serbi.ula.ve/ematres/	
Glosarios					
No.	Título	Responsable	Año	Ubicación	
1	Glosario de Términos Siderúrgicos	Grupo CAP	1995	Biblioteca de la Universidad de Moa	

2	Glosario Mecánica Industrial		2019	https://glosarios.servidor-alicante.com/mecanica-industrial
3	Glosario de Términos mecánicos	DEMAC	2013	http://demacmotor.com/glosario-de-terminos-mecanicos-c/
4	Glosario de Energía	Empresaseléctricas A.G		https://www.electricas.cl/educacion-en-energia/glosario-de-energia/
5	Glosario de términos de energía	Eligetuenegía		https://eligetuenegia.com/glosario/
6	Glosario utilizado en minería y metalurgia	ACADEMIA		https://www.academia.edu/37105981/Glosario_de_terminos_usados_en_mineria_y_metalurgia
7	Tesouro en línea	Biblioteca Dr. Jorge Villalobos Padilla, S.J.	2014	http://quijote.biblio.iteso.mx/catia/tesouro/default.htm
Otras recursos consultados				
No.	Nombre	Ubicación		
1	Monografías	https://www.monografias.com/		
2	Flexicon	http://www.flexicon.es/Equipos-y-Sistemas-de-Manejo-de-Materiales-a-Granel/index.html		
4	Cimsa	http://www.cimsaweb.com.ar/index		
5	SCRIBD	https://es.scribd.com/docs		
6	Google	https://www.google.com/		
Planes de estudio				
No.	Nombre	Especialidad		

1	Plan de estudios E	Ing. Eléctrica
2	Plan de estudios E	Ing. Mecánica
3	Plan de estudios D	Ing. Mecánica
4	Plan de estudios E (curso diurno)	Ing. Mecánica

1.f) Lineamientos y reglas establecidas para presentar los términos.

1. Forma del término

Utilizar siempre que sea posible, los sustantivos. No utilizar verbos.

2. Número del término

Utilizar el singular para términos que expresen conceptos que no puedan contarse por unidades.

Utilizar el plural para términos que expresan conceptos que puedan contarse por unidades.

Tabla 4. Lineamientos para presentar los términos.

<i>Términos</i>	<i>Usar singular cuando:</i>	<i>Usar plural cuando:</i>
Términos que denotan materiales, herramientas, equipos y piezas	Sea un término específico (TE). Ejemplo: Acero inoxidable	Sea un término genérico (TG). Ejemplo: Aceros
Términos que denotan propiedades, características y parámetros.	Viscosidad Elasticidad Temperatura Combustión	Propiedades mecánicas Propiedades físicas Propiedades químicas
Términos que denotan procesos.	Beneficio de minerales Cribado Lixiviación Oxidación Sedimentación	-

	Contaminación Corrosión	
Términos que denotan disciplinas, materias, campos del saber.	Hidráulica Química	-
Términos que denotan nombres propios de métodos, técnicas, procedimientos.	Ley de Kirchhoff Ley de Ohm	-
Término que denota nombres de sustancias.	Sea un término específico: Aceite RIDGID Etanol	Sea un término genérico Aceites
Términos que denotan agentes.	-	Especialistas Ingenieros Estudiantes
Términos que denotan tipo de investigación.	-	Estudios críticos Bibliografías Ensayos
Términos que denotan fenómenos.	-	Explosiones

3. Forma de entrada

Se utilizará la entrada directa.

4. Se evitará la utilización de signos de puntuación, y abreviaturas.

5. En el proceso de indización se trabajará con el resumen de los documentos, en el caso que así lo requiera se utilizará el documento completo.

6. Se trabajará con un vocabulario específico, siempre y cuando los términos a utilizar se manejen con bastante frecuencia y la práctica lo demuestre imprescindible.

7. Los descriptores se presentarán utilizando la mayúscula inicial.

8. Los términos no autorizados se pondrán en letra cursiva.
9. El tesauruso aceptará el uso de calificadores, para aclarar el alcance de los términos, solo se permite el uso de corchetes al final del término como forma aclaratoria.
10. Se utilizará el guion solo en nombres oficiales, nombres propios o fórmulas químicas que contienen como parte integrante del término el guion.

Segunda Fase

2.a) Someter a revisión cada palabra clave capturada.

Las palabras claves capturadas se sometieron a un proceso de revisión por parte de un grupo de especialistas, integrado este por:

Tabla 5. Especialistas que revisaron la terminología del tesauruso.

Nombre y apellidos	Categoría científica	Departamento docente	Años de experiencia
Yordan Guerras Rojas	MScen Ciencias	Eléctrica	16
Reineris Montero Laurencio	Dr. en Ciencias	CEETAM	20
Yosbel Guerra González	MSc. en Ciencias	Metalurgia-Química	6
Roger Samuel Almenares Reyes	Dr. en Ciencias	Metalurgia-Química	13
Tomás Fernández Columbré	Dr. En Ciencias	Mecánica	29
Isnel Rodríguez González	Dr. en Ciencias	Mecánica	26

A este grupo de trabajo se le solicitó colaboración para revisar la terminología de base utilizada, para precisar los términos, crear las relaciones entre ellos y emitir criterios y sugerencias. (**Ver anexo1**)

Se utilizó además el criterio de las especialistas que realizan el proceso de la indización, para conocer la utilidad de los términos desde el punto de vista de la indización, la recuperación de la información y los intereses informativos de la comunidad usuaria del tesoro.

A partir de este análisis surgieron nuevos términos y otros se modificaron o eliminaron.

Tabla 6

Tercera Fase

3.a) Elaboración de artículos léxicos.

3.a.1) En este paso se crearon todas las relaciones entre los términos: relaciones jerárquicas, de equivalencia, asociativas y se redactaron las notas de alcance convenientes. Para establecer las relaciones entre los términos fue necesario utilizar herramientas léxicas y el criterio de los expertos.

Tabla 7. Tabla de referencia.

Designación en español.	Significado de la referencia	Tipo de referencia.
TG	Término genérico	Referencia del descriptor al descriptor de nivel superior.
TR	Término relacionante	Referencia del descriptor al descriptor de nivel inferior.
TE	Término específico	Referencia del descriptor al descriptor de nivel inferior.
NA	Nota de alcance	Precisión del término
USE	Término preferido	Referencia del no descriptor al descriptor.
UP	Usado por	Referencia del descriptor al descriptor sinónimo o al descriptor.

3.a.2) Las referencias señalan las relaciones de una unidad léxica con otras, y son el resultado de las operaciones siguientes:

- Eliminación de la homonimia de las unidades léxicas.
- Establecimiento de las relaciones de sinonimia
- Establecimiento de las relaciones jerárquicas y asociativas de los descriptores.

Para eliminar la homonimia de las unidades léxicas, se asignan a estos calificadores los cuales son parte inseparable de la unidad léxica. Ejemplo:

Árboles [mecánica]

Cuando se hace necesario se redactan notas de alcance, (NA) las cuales no forman parte inseparable de la unidad léxica. Estas se usan para:

- precisar el significado del término.
- limitar el campo de aplicación.
- destacar un significado determinado del término, si el término se emplea en diferentes disciplinas.
- incluir fechas de inclusión y exclusión de los términos y registros de los cambios.

Ejemplo: Árboles [mecánica]

NA *Elemento para transmisión del movimiento giratorio*

3.a.3) Selección de los descriptores.

Para seleccionar los descriptores en la clase de sinonimia debe seleccionarse una unidad léxica como representante de la misma.

3.a.3.1) Los criterios principales para seleccionar una unidad léxica como descriptor son:

- que exprese exhaustivamente el significado de la clase de sinonimia.
- que la unidad léxica sea clara y precisa.
- que tenga una frecuencia de utilización en la búsqueda e indización de los documentos.

3.a.3.2) En calidad de no descriptores se utilizan:

- cualquier unidad léxica que integre la clase de sinonimia y que no es representativa de la misma.
- la unidad léxica cuyo significado se representa en la búsqueda informativa como intersección de significado de otras unidades léxicas que funcionan como descriptores.
- la unidad léxica cuyo significado es representado en la búsqueda informativa como la unión de significados de otras unidades léxicas incompatibles en un modelo de búsqueda.

3.a.3.3) Los artículos léxicos de los no descriptores deben de tener referencia a los descriptores que lo sustituyen. Por ejemplo:

No descriptor Ahorro energético: energética

descriptor: USE Ahorro de energía

3.a.3.4) En el tesauro se debe señalar la existencia de la relación jerárquica de los descriptores indicando las relaciones establecidas en la tabla 8.

Ejemplo:

TG Bombas

TE Bomba centrífuga

3.a.3.5) La relación parte-todo se establece entre dos descriptores cuando el descriptor de nivel inferior es componente del objeto designado como descriptor del nivel superior.

Ejemplo:

Automóviles

Carrocería

3.a.3.6) Relaciones asociativas: Se permiten incluir en estas relaciones los términos capaces de asociar ideas y relaciones entre ellos.

Ejemplo:

descriptor: Aceites lubricantes

TR Viscosidad

3.a.3.7) Revisar con los especialistas el artículo léxico ya establecido.

Se verificó que los términos estuvieran ubicados en los subcampos establecidos, se cumplieran las relaciones recíprocas entre los descriptores, el uso correcto de los calificadores y la exactitud y veracidad de las notas de alcance.

Cuarta Fase

4.a) Automatización en la elaboración del tesoro.

Existen diversas herramientas para la gestión de tesauros, en la mayoría de los casos son fáciles de implementar y de usar, algunas de ellas son comerciales y otras son libres, pero a la hora de seleccionar alguna, es importante evaluar previamente cuál es la herramienta tecnológica que más se adapta a las necesidades, teniendo siempre presente lo que se busca o persigue con su explotación.

Para efectos del presente trabajo se empleó el software TemaTres, el cual es una aplicación Web libre para la gestión de lenguajes documentales.

Se encuentra orientado especialmente al desarrollo de tesauros jerárquicos, además puede utilizarse para desarrollar estructuras de navegación Web, o como complemento articular con un gestor de contenidos, bibliotecas digitales, o en una biblioteca tradicional para gestionar los lenguajes documentales en uso.

Su sitio oficial es <http://www.vocabularyserver.com/index.html> desde donde se puede descargar el software.

La utilización de este software permite automatizar algunas etapas en la elaboración del tesoro. Por ejemplo:

- análisis estadístico
- corrección de artículos léxicos
- alfabetización del listado
- control de las referencias
- composición de índices
- presentación del tesoro en determinado formato

4.b) Presentación del tesoro especializado en Ciencia de los Materiales. **(Ver anexo 2)**

TESAURO DE CIENCIAS DE LOS MATERIALES

Introducción:

El Tesoro de Ciencia de los Materiales compila un total de 1603 términos, donde existen 2044 relaciones entre los términos y 120 términos equivalentes, en la edición de junio 2019. La creación del Tesoro se ha basado en los términos presentes en el Repositorio Institucional Nínive de la Universidad de Moa Dr. Antonio Núñez Jiménez, generados a partir del proceso de indización de la colección documental de la facultad de Metalurgia-Electromecánica.

El tesoro se encuentra en un período de gestación y desarrollo, por lo que su actualización y mantenimiento será responsabilidad de la biblioteca de la Universidad de Moa Dr. Antonio Núñez Jiménez.

Es un vocabulario monolingüe que se aplica a la indización y recuperación de la información almacenada.

Para la inclusión de nuevos términos se ha tomado en cuenta los procedimientos establecidos en la literatura especializada de las Ciencias de la Información y la temática abordada. Para la eliminación de términos se ha tenido en cuenta si la frecuencia de uso de los mismos es nula o poco frecuente. De igual manera se han analizado las relaciones jerárquicas, de equivalencia, asociativas y confección de notas de alcance.

Para la elaboración del tesoro se utilizó como herramienta bibliotecológica el Sistema de Clasificación Decimal Dewey y herramientas léxicas especializadas.

El tesoro está dirigido a estudiantes, profesores e investigadores de la comunidad universitaria de la Universidad de Moa Dr. Antonio Núñez Jiménez. Además de aquellas personas e instituciones interesadas en consultar dicha información.

Nota de uso del tesoro.

El presente tesoro consta de dos partes:

- Sección Semántica
- Sección Alfabética

Sección Semántica

Muestra el orden de los descriptores por temas o campos semánticos.

Ingeniería eléctrica

Ingeniería mecánica

Ingeniería metalúrgica y química

Tecnología ambiental

Estas categorías se encuentran a su vez subdivididas por áreas o microtesauros más específicos. Se ofrece en conjunto los distintos descriptores que se asocian bajo determinada área temática, lo que permite la elección del término acertado.

Sección Alfabética

Muestra el listado alfabético de los términos y sus relaciones. Los símbolos usados para representar en el tesauruso son:

NA Nota de alcance

Precisa el sentido del descriptor para fines de indización.

Calificadores []

Precisa el alcance del término en el tesauruso. Elimina la homonimia.

UP Usado por

Señala los términos sinónimos o no descriptores; que no son válidos para la indización.

TG Término genérico

Señala el término más amplio o genérico al que pertenece el descriptor.

TE Término específico

Señala los términos derivados (tipos o clases) del descriptor base.

TR Término relacionado

Señala los términos relacionados semánticamente con el descriptor base.

Ejemplo:

Baterías

TG Electroquímica aplicada

TE Baterías de iones de litio

TR Corriente directa

Sistemas fotovoltaicos autónomos

Cuando un mismo concepto se puede expresar mediante dos o más sinónimos, uno de ellos tiene que seleccionarse como término preferente o descriptor. Los demás pasan a ser no descriptores.

Ejemplo: *Abasto de agua* **USE** Abastecimiento de agua

El tesoro ha sido elaborado por la estudiante Ydelsi Vielza Caraballo, en opción al título de Lic. en Ciencias de la Información bajo la tutoría de la Lic. Niurka de la Vara Garrido especialista principal en gestión de la información de la biblioteca de la Universidad de Moa y al Dr. en ciencia José Luis Montero O'farrill, director del Centro de Tecnología Educativa de la Universidad de Moa.

Se agradece la colaboración de:

MSc. Yordan Guerra Rojas

Dr. Reineris Montero Laurencio

MSc. Yosbel Guerra González

Dr. Roger Samuel Almenares Reyes

Dr. Tomás Fernández Columbré

Dr. Isnel Rodríguez González

Quienes aportaron sus conocimientos y experiencias en las materias de mecánica, eléctrica y metalurgia-química, haciendo posible este producto.

Tesauro Ciencia de los Materiales

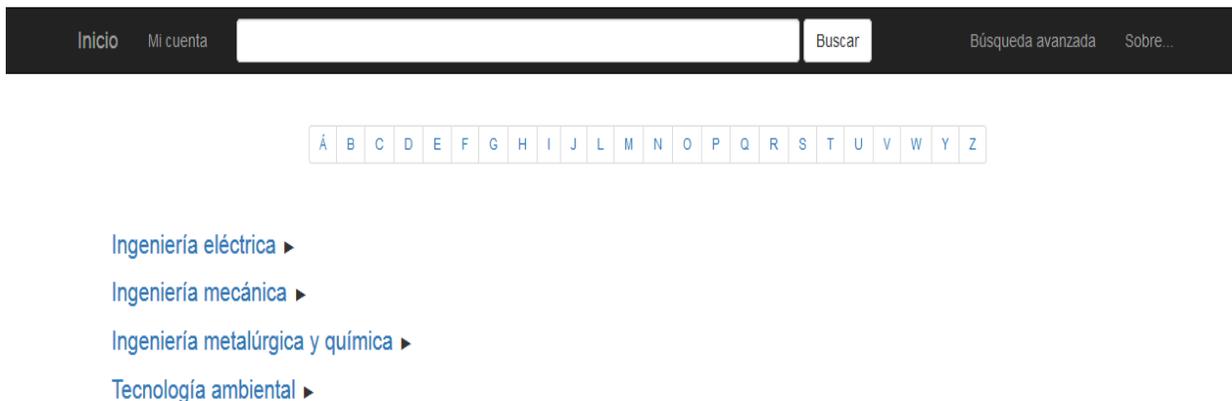


Fig. 3. Vista de la página inicial del tesauro en línea.

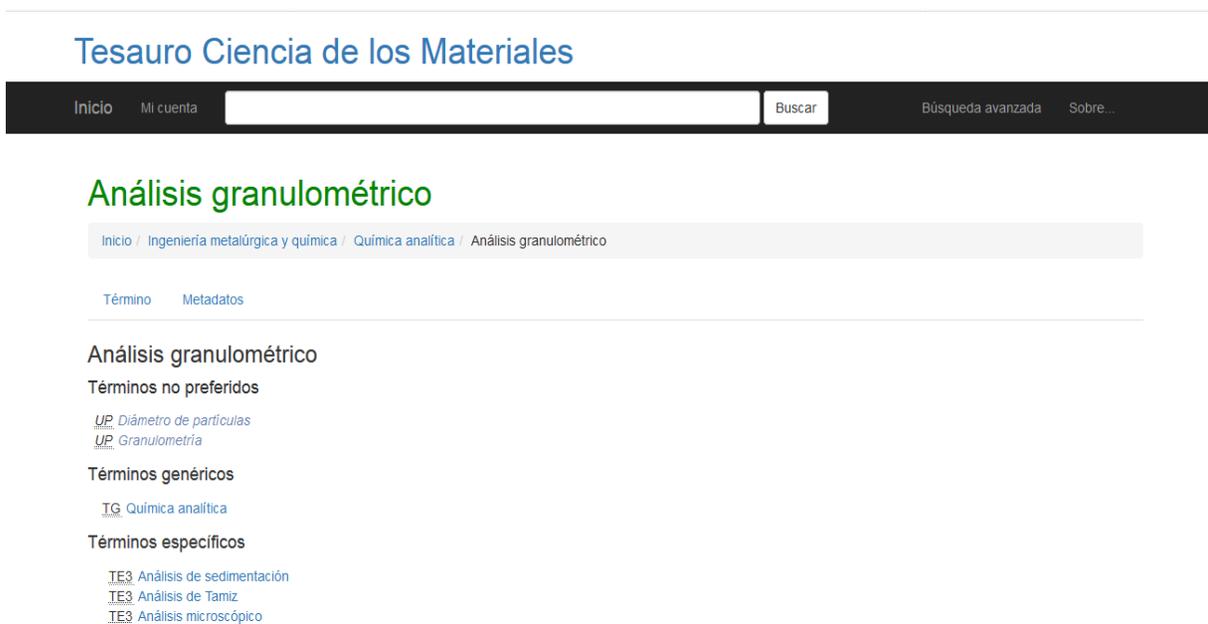
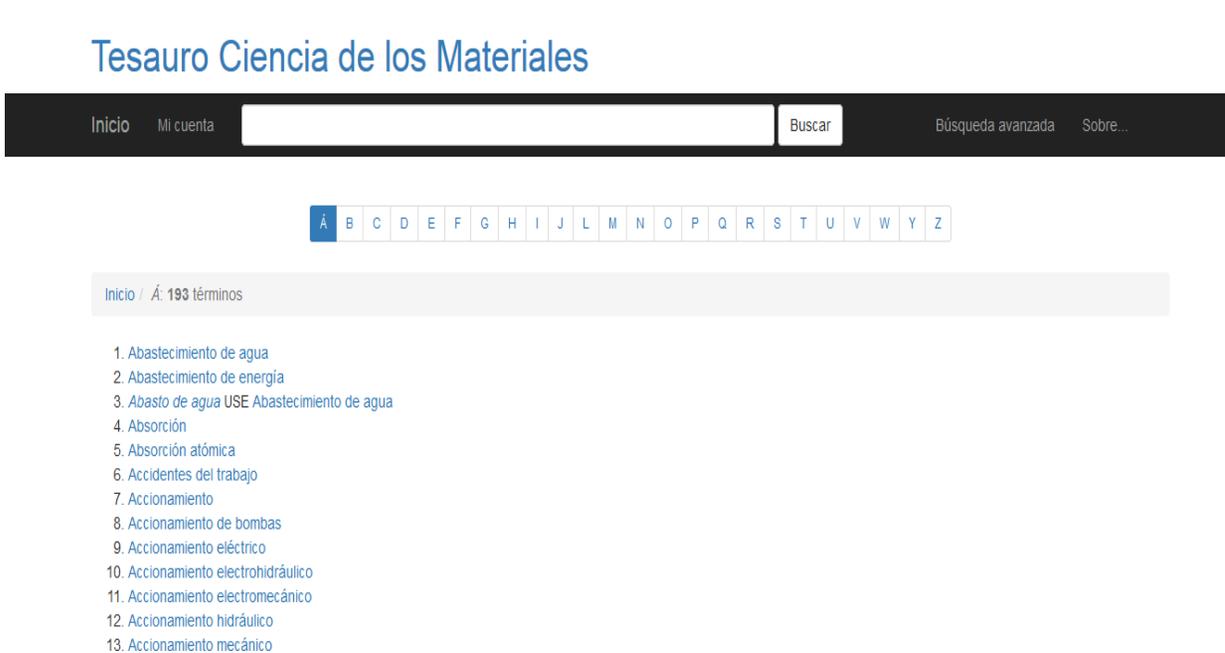


Fig.4. Vista de la presentación jerárquica del tesauro.



The screenshot shows the 'Tesauro Ciencia de los Materiales' website. At the top, there is a navigation bar with 'Inicio', 'Mi cuenta', a search input field, and a 'Buscar' button. To the right of the search bar are links for 'Búsqueda avanzada' and 'Sobre...'. Below the navigation bar is an alphabetical index from 'A' to 'Z', with 'A' highlighted. Underneath the index, a grey bar indicates 'Inicio / A: 193 términos'. A list of terms is displayed, starting with '1. Abastecimiento de agua' and ending with '13. Accionamiento mecánico'.

Fig. 5. Vista de la presentación alfabética del tesauro.

4.e) Manual de usuario de TemaTres 2.0.

TemaTres es una herramienta Web para la gestión y explotación de vocabularios controlados, tesauros, taxonomías y otros modelos de representación formal del conocimiento.

TemaTres permite implementar modelos distribuidos de gestión garantizando condiciones de consistencia e integridad de datos y relaciones entre términos. Dispone de funcionalidades especialmente orientadas a ofrecer datos de trazabilidad y control de calidad en el contexto de un vocabulario controlado. Permite a su vez la representación de vocabularios a través de todos los estándares de metadatos disponibles en el campo de la gestión del conocimiento.

❖ Instalación y configuración

Requerimientos

Los requerimientos de TemaTres son los siguientes:

- Un servidor web. La mayoría de los usuarios utiliza Apache, pero TemaTres debe funcionar bien en cualquier servidor web que soporte PHP, como el IIS (Internet Information Server) de las plataformas Windows.
- Una instalación de PHP en funcionamiento (versión 4.3.0 o posterior). PHP 5 está soportado.
- Un servidor de bases de datos. Por defecto TemaTres utiliza MySQL, pero puede utilizar una variedad de motores de bases de datos (Postgres, Oracle MS server, MySQL). Detalles acerca de los servidores de bases de datos soportados.

Los tres pasos previos al proceso de instalación:

Paso 1) Descomprima TemaTres en el directorio de acceso web de su servidor (ej: /var/www/tematres/ o c:\xampp\htdocs\tematres\)

Paso 2) Configure la conexión con la base de datos en el archivo db.tematres.php disponible en el directorio por defecto del vocabulario controlado /vocab/ (ej:

/var/www/tematres/vocab/db.tematres.php o
c:\xampp\htdocs\tematres\vocab\db.tematres.php.

Los datos a configurar son los siguientes:

Selección del tipo de servidor de bases de datos a utilizar. Si se deja en blanco, utilizará MySQL.

```
$DBCFG["DBdriver"] = "";
```

Dirección IP o nombre del servidor de bases de datos, por ejemplo, localhost

```
$DBCFG["Server"] = "localhost";
```

Nombre de la base de datos

```
$DBCFG["DBName"] = "tematres";
```

Nombre de usuario para conectarse con la base de datos

```
$DBCFG["DBLogin"] = "root";
```

Clave de usuario para conectarse con la base de datos

```
$DBCFG["DBPass"] = "";
```

Prefijo para las tablas de este vocabulario controlados

```
$DBCFG["DBprefix"] = "lc_";
```

Paso 3) Vaya con su navegador web al directorio de instalación de TemaTres y siga las instrucciones de instalación.

Fig.6. Vista de la instalación del software Tema Tres

El usuario ingresado será el administrador del vocabulario.

Verificar que el correo electrónico está correctamente escrito.

Una vez configurado podrá ingresar a su vocabulario. Cliquee en Mi Cuenta e ingrese los datos solicitados.

El siguiente paso es ingresar en Menú / Administración / Configuración:

Haciendo clic sobre el nombre del vocabulario podrá definir algunos datos de configuración. Los más importantes son: título, autor, cobertura y el correo electrónico.

Tesoro Ciencia de los Materiales

Inicio Menú ▾ Agregar término Buscar Búsqueda avanzada Sobre...

Configuración

Título	Autor	Tipo
Tesoro Ciencia de los Materiales / es	Ydelsi	vocabulario

Fig.7. Vista del producto de la configuración del software.**4.f) Actualización del software TemaTres.**

Desde la versión 1.2 de TemaTres es posible utilizar rutinas automatizadas para la actualización desde versiones anteriores. Las opciones de actualización disponibles son:

- desde la versión 1.0 hacia la versión actual
- desde la versión 1.1 hacia la versión actual

Para actualizar la instalación de TemaTres debe:

1. Por razones de seguridad, hacer un respaldo de la base de datos en uso con su antigua versión de TemaTres.
2. Descargar e instalar la versión más reciente de TemaTres
3. Ingresar como administrador (utilizando sus anteriores nombres de usuario y clave)
4. Ir al Menú superior → Opción Administración → Sub-opción Mantenimiento de la Base de datos
5. Seleccionar la actualización a realizar (desde la versión 1.0 o 1.1).

4.g) Actualización del tesoro.

El crecimiento de la información, la diversidad de temáticas de investigación y la aparición de nuevos términos sobre el área de la Ciencia de los Materiales en el Repositorio Institucional Nínive será la condición que posibilitará la actualización y ajustes en el Tesoro.

La periodicidad de la revisión será anual.

Los lineamientos establecidos con esta investigación y el tesoro creado implicarán correcciones en la carga de los metadatos de materia asignados a la colección documental de la Facultad de Metalurgia Electromecánica, lo que deberá actualizarse los registros de los documentos en el Repositorio.

Quinta fase

5.a) Comunicar y difundir el producto terminológico confeccionado.

Una vez concluida la etapa de elaboración del tesoro, se puede difundir el tesoro como un producto terminológico:

Tesoro en línea: accesibles para usuarios en la Web con disponibilidad de consultarlo, permite:

Publicar y descargar en varios formatos.

Uso directo en centros de documentación, bases de datos, páginas web y repositorios.

Tesoro en formato físico: los documentos y las impresiones en papel se pueden aprovechar con mayor facilidad, ya que se trata de un formato sobre el que cualquiera puede trabajar directamente sin necesidad de disponer de ningún tipo de medio concreto. Puede ser consultado sin necesidad de un aparato tecnológico para su reproducción. Los documentos e impresiones quedan a salvo de los peligros del mundo digital, una realidad que ha logrado dar un importante impulso al papel en plena era digital.

CONCLUSIONES

- El tesoro es una herramienta necesaria para la normalización de los términos, siendo identificado como un instrumento de control terminológico que contribuye a la recuperación de la información de forma precisa y efectiva.
- El tesoro de Ciencia de los Materiales es un lenguaje de indización que permite identificar los temas tratados en los documentos sobre este dominio temático en la Universidad de Moa.
- El tesoro de Ciencia de los Materiales refleja las expresiones utilizadas por los autores de los documentos y los usuarios en sus formulaciones de búsqueda de información.
- El método de comité el método empírico permitieron generar el tesoro con un total de 1603 de términos normalizados y 20 lineamientos de indización.
- El software utilizado para la generación del tesoro, posibilita desarrollar la estructura necesaria entre los términos para su implementación como herramienta de trabajo.
- El tesoro especializado en el área de la Ciencia de los Materiales propone a la comunidad científica de la Universidad de Moa Dr. Antonio Núñez Jiménez un conjunto estructurado de términos que unifican y enriquecen el núcleo de términos utilizados actualmente, para el procesamiento y la recuperación de la información.

RECOMENDACIONES

1. Complementar esta investigación con la implementación del Tesauro de Ciencia de los Materiales de la Universidad de Moa Dr. Antonio Núñez Jiménez.
2. Mantener y actualizar el Tesauro de Ciencia de los Materiales para que pueda ser utilizado como obra de trabajo o de consulta para los indizadores y usuarios del repositorio.

REFERENCIA BIBLIOGRÁFICA

- Albuquerque, P. C. d. (2018). *Soporte de vocabularios controlados y autoridades en repositorios digitales*. Facultad de Informática. Disponible en: <http://sedici.unlp.edu.ar/bitstream/handle/10915/69754/Tesis.pdf?sequence=1>
- Almazán, M. B. (2012). *Herramientas para la interoperabilidad y normalización de datos en RI*. Facultad de Informática. Disponible en: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=2ahUKEwis_4vb6_DiAhUkwlkKHZ2bAtEQFjABegQIARAC&url=http%3A%2F%2Fcatalogo.info.unlp.edu.ar%2Fmeran%2FgetDocument.pl%3Fid%3D384&usg=AOvVaw1LJ2DlqVcoZR96klUMu8X1
- Barité, M. (2014). El control de vocabulario en la era digital: revisión conceptual. *Scire*, 20(1). Disponible en: <https://www.ibersid.eu/ojs/index.php/scire/article/view/4196/3766>
- Bernal, I. (2016). *Uso de Vocabularios Controlados en Repositorios. La experiencia de DIGITAL.CSIC: CSIC - Unidad de Recursos de Información Científica para la Investigación (URICI)*. Disponible en: http://digital.csic.es/bitstream/10261/140742/1/Bernal_DIGITALCSIC_MECD.pdf
- Biblioteca Dr. Jorge Villalobos Padilla, S.J. (2014). Disponible en: <http://quijote.biblio.iteso.mx/catia/tesauro/default.htm#>
- Cano Inclán, A., de Dios Arias, R. A., García García, O., & Cuesta Rodríguez, F. (2015). Los repositorios institucionales: situación actual a nivel internacional, latinoamericano y en Cuba. *Revista Cubana de Información en Ciencias de la Salud*, 26, 0-0. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2307-21132015000400002&nrm=iso
- Cañedo Andalia, R., y Small Chapman, M. C. (2011). Elementos sobre indización y búsqueda de la información por medio de vocabularios controlados en bases de datos biomédicas. *Revista Cubana de Información en Ciencias de la Salud (ACIMED)*, 22(2), 142-154. Disponible en: <https://www.medigraphic.com/pdfs/acimed/aci-2011/aci112e.pdf>

CIMSA. Disponible en: <http://www.cimsaweb.com.ar/index.html>

Flexicon. (2015). Disponible en: <http://www.flexicon.es/>

The free dictionary. (2018). from <https://es.thefreedictionary.com/>

García Walters, J. (2004). *Tesaurus de Ingeniería Sanitaria y Ambiental*. Disponible en: <http://www.bvsde.paho.org/bvsair/e/manuales/tesa/teses.pdf>

Hernández Quintana, A. R. (2004). *Indización y Resumen*. La Habana: Editorial Félix Varela.

Diccionario Español de Ingeniería. Disponible en: <http://diccionario.raing.es/es/page/modo-de-citar>

López Veulens, S. (1983). *Indización II*. La Habana: Ministerio de Educación Superior.

Machado Rivero, M. O. (2017). *Política para el desarrollo del sistema de repositorios digitales en la red de bibliotecas universitarias del Ministerio de Educación Superior*. La Habana. Disponible en: <http://cep.ismm.edu.cu:8080/xmlui/handle/123456789/1338>.

Malavassi Aguilar, F. A. (2007). *Propuesta para la construcción de un Tesaurus Especializado en Arquitectura*. (Tesis de Diploma), Universidad de Costa Rica, Costa Rica. Disponible en: <http://repositorio.sibdi.ucr.ac.cr:8080/jspui/bitstream/123456789/253/1/28444%20Tomo%20I.pdf>

Martín Gavilán, C. (2009). *Temas de Biblioteconomía: Lenguajes documentales. Principales tipos de clasificación. Encabezamientos de materia, descriptores y tesaurus*. Disponible en: <http://eprints.rclis.org/14817/1/lendoc.pdf>

Martínez Ferreras, D. (2017). *Los tesaurus*. Disponible en: https://www.formacion.andaluciaesdigital.es/c/document_library/get_file?uuid=d1f4b8e7-702e-48d2-94b4-88846907cf26&groupId=20195

Monografías PLUS+. Disponible en: <https://www.monografias.com/>

MotorGiga. from <https://diccionario.motorgiga.com/>

- Müller S., F. Diccionario Técnico de Mecánica de Automoviles. 1ra ed. Disponible en: <https://www.bcn.cl/obtienearchivo?id=documentos/10221.1/61186/1/207689.pdf&origen=BDigita>
- NC 39-14 1983. *Elaboración del Tesauro Unilingüe de Búsqueda Informativa*. (1983). La Habana: Comité Estatal de Normalización.
- Osés, M. A. (2017). *Manual de TemaTres. La consulta y gestión de un vocabulario controlado*. Disponible en: <http://vocabularios.saij.gob.ar/portalthes/descargas/ManualTematresSAIJ2017.pdf>
- Pinzón Otálora, Ó. J. *Construcción de un tesauro especializado en comunicación social para el contexto colombiano*. Disponible en: http://repository.lasalle.edu.co/bitstream/handle/10185/21581/33052213_2017.pdf?sequence=1
- Ramírez Polo, L. (2012). Los lenguajes controlados y la documentación técnica: mejorando la traducibilidad. *Tradumàtica*(10), 0192-0204. Disponible en: https://ddd.uab.cat/pub/tradumatica/tradumatica_a2012n10/tradumatica_a2012n10p192.pdf
- Redondo Quintela, F. (2006). Diccionario de Ingeniería Eléctrica. from <http://electricidad.usal.es/Principal/Circuitos/Diccionario/Inicio.php>
- Rodríguez Noguera, M. (2006). *Propuesta preliminar de diseño de un Tesauro de la Música Cubana para uso de instituciones de información* Universidad de La Habana, La Habana. Disponible en: <http://10.28.1.80:9090/cgi-bin/utf8/wxis/iah/scripts/>
- Rubricador del Sistema Nacional de Información Científica y Técnica* (1988). (2 ed. Vol. III t.). La Habana: Instituto de Documentación e Información Científica y Técnica.
- SCRIBD. from <https://es.scribd.com/docs>
- Soler Monreal, C., & Gil Leiva, I. (2010). Posibilidades y límites de los tesauros frente a otros sistemas de organización del conocimiento: folksonomías, taxonomías y ontologías. *Revista Interamericana de Bibliotecología*, 33(2), 361-377. Disponible en :

<https://riunet.upv.es/bitstream/handle/10251/6341/tesisUPV3096.pdf?sequence=1&isAllowed=y>

Soler Monreal, M. C. (2009). *Evaluación de vocabularios controlados en la indización de documentos mediante índices de consistencia entre indizadores*. Disponible en:

<https://riunet.upv.es/bitstream/handle/10251/6341/tesisUPV3096.pdf?sequence=1&isAllowed=y>

Tesauro EUROVOC (2006). Disponible en:

http://www.bizkaia.eus/descargar_documento.asp?url=kultura%2Feurovoc%2Fidb%2Fpdfs%2Fpermut_es.pdf¶m=1

Tesauro de Materiales de Serbiula. (2015). Disponible en:

<http://www.serbi.ula.ve/tematres/vocab/index.php>

Tesauros del Patrimonio Cultural de España. (2019). Disponible en:

<http://tesauros.mecd.es/tesauros/materias/1188762.html>

The free dictionary. (2018). Disponible en: <https://es.thefreedictionary.com/>

Texier, J., De Giusti, M. R., Oviedo, N., Lira, A. J., & Villarreal, G. L. (2013). *La representación de recursos en los repositorios institucionales*. Paper presented at the Conferencia de Bibliotecas y Repositorios Digitales de América Latina y Simposio Internacional de Bibliotecas Digitales (Costa Rica, 2013). Disponible en: https://digital.cic.gba.gob.ar/bitstream/handle/11746/3549/11746_3549.pdf-PDFA.pdf?sequence=1&isAllowed=y

Vázquez López, J. (1974). *Diccionario Enciclopédico de la Siderurgia* (1 ed.). España: Urmo Ediciones.

Vizcaya Alonso, D. (1999). Gerencia de los lenguajes documentarios. *Ciencias de la Información*, 30(2), 51-60. Disponible en:

<http://cinfo.idict.cu/index.php/cinfo/article/download/292/291>

Yedid, N. (2014). *Modelos de indización temática utilizados en repositorios digitales de acceso abierto de Argentina*. Disponible en:

https://www.bn.gov.ar/resources/conferences/pdfs/YedidN_Ponencia.pdf

ANEXOS

Anexo I:

Universidad de Moa Dr. Antonio Núñez Jiménez

Facultad Metalurgia-Electromecánica

Especialidad: _____

Encuesta para obtener criterios a utilizar en la confección y desarrollo del Tesauro en Ciencia de los Materiales.

Departamento docente: _____ Cargo: _____

Profesor: _____

Categoría docente: _____

Grado científico: _____

Años de experiencia en la actividad: _____

Con vista a desarrollar un Tesauro en el dominio temático de la Ciencia de los Materiales a ser utilizado en el procesamiento, búsqueda y acceso a la información disponible en el Repositorio Institucional de la Universidad, se desarrolla un trabajo de diploma en la especialidad de Ciencias de la Información, para el cual solicitamos su colaboración en la revisión del vocabulario¹ que se pone a su consideración.

Específicamente se precisa de usted:

- Sugerencia de *términos sinónimos* a los que se brindan.
- Sugerencia de *términos relacionados*.
- Sugerencia de *términos genéricos*.
- Sugerencia de *términos específicos*
- Sugerencia de *notas aclaratorias* en algún término que así lo requiera.

Gracias por su valiosa colaboración.

¹ Vocabulario generado durante el procesamiento analítico sintético que se le realiza a los documentos que van a formar parte del Repositorio Institucional, considerando las palabras claves utilizadas por el autor de los documentos y otras asignadas por las especialistas del Centro de Información.

Anexo2:

Ejemplo para indizar con el tesauro de Ciencia de los Materiales.

Título de la tesis: Diagnóstico de los parámetros de explotación de bombas centrífugas de mediana capacidad en presencia de régimen cavitacional.

Autor: Reyes de la Cruz, Jorge Luis

Comunidad: Metalurgia-Electromecánica

Tesauro utilizado: Tesauro Ciencia de los Materiales

Microtesauro: Ingeniería mecánica

Resumen: En la investigación se fundamenta de forma experimental los niveles de vibraciones en correspondencia con las alturas de succión, las etapas del régimen cavitacional y se plantea el procedimiento para el diagnóstico de instalaciones con bombas centrífugas afectadas por este fenómeno. Se define la presencia de la cavitación a partir del conocimiento de los factores hidráulicos que influyen sobre la carga neta de succión de la instalación, además las ecuaciones que definen la presencia del fenómeno de cavitación, la comparación de los valores reportados por el fabricante con los reales, y los niveles de vibraciones en los que la instalación debe operar, para garantizar el dominio de los parámetros técnicos y económicos de la instalación.

1. Análisis de contenido.

- Lectura técnica del documento y la extracción de los conceptos o términos relevantes. Se le prestó atención al resumen y al título del documento, del cual se extrajeron las siguientes palabras claves:

Bombas centrífugas

Cavitación

Vibraciones

- Se consultaron las palabras claves asignadas con el tesauro.

- Bombas centrífugas: se encuentra este término en el tesoro, pero en singular, se asigna como término de indexación. Donde se asigna además con carácter obligatorio su descriptor genérico (TG)- Bombas, donde aparece el término usado por (UP)- Bomba rotodinámica.

- Cavitación: aparece como término de entrada, por tanto se asigna como término de indexación.

- Vibraciones: aparece el término, por lo que se asigna como término de indexación.

Notas: En caso de no aparecer en la parte léxica o cuerpo central del microtesoro el término buscado, se analiza si el término es un nombre propio, si es un nombre propio, se asigna como descriptor.

De no ser un nombre propio se procede a buscar posibles sinónimos del término en el tesoro. Si aparece algún término sinónimo con una referencia de USE, entonces se utiliza el descriptor autorizado.

De no aparecer ningún sinónimo del término se procede a estudiar la posibilidad de incluir el término en el tesoro como candidato a descriptor. Si se llega a la conclusión que no debe ser incluido en el tesoro, no se asigna como término descriptor.

Conclusión: La indexación que debe tener el documento seleccionado, ha de comprender como descriptores:

Bombas, Bomba centrífuga, Cavitación y Vibraciones.

ⁱ Vocabulario generado durante el procesamiento analítico sintético que se le realiza a los documentos que van a formar parte del Repositorio Institucional, considerando las palabras claves utilizadas por el autor de los documentos y otras asignadas por las especialistas del Centro de Información.