

**CONSTRUCCIÓN DE UN TESAURO ESPECIALIZADO EN HIDROCARBUROS**

**JACKSON ANDRES MARTÍNEZ TRIANA**

**UNIVERSIDAD DE LA SALLE  
FACULTAD DE CIENCIAS ECONÓMICAS Y SOCIALES  
PROGRAMA DE SISTEMAS DE INFORMACIÓN Y DOCUMENTACIÓN  
BOGOTÁ  
2015**

**CONSTRUCCIÓN DE UN TESAURO ESPECIALIZADO EN HIDROCARBUROS**

**JACKSON ANDRES MARTÍNEZ TRIANA**

**Trabajo de grado para optar por el título de  
Profesional en Sistemas de Información, Bibliotecología y Archivística**

**Director**

**EDWIN ALONSO SOSSA CRUZ**

**Bibliotecólogo**

**UNIVERSIDAD DE LA SALLE  
FACULTAD DE CIENCIAS ECONÓMICAS Y SOCIALES  
PROGRAMA DE SISTEMAS DE INFORMACIÓN Y DOCUMENTACIÓN  
BOGOTÁ  
2015**

Nota de aceptación:

---

**Dr. NELSON PULIDO**  
Director del Programa de Sistemas de Información y Documentación

---

**EDWIN ALONSO SOSSA CRUZ**  
Director del Trabajo de Grado

---

**Jurado**

---

**Jurado**

Bogotá D.C., 08 de julio de 2015.

## **AGRADECIMIENTOS**

En primer lugar deseo expresar mi agradecimiento al director de esta tesis, profesor. Edwin Alonso Sossa Cruz, por su esfuerzo y dedicación, quien con sus conocimientos, su experiencia, su paciencia y su motivación ha logrado en mí que pueda terminar mis estudios con éxito.

Del mismo modo agradezco a mis padres, hermano y a Dios por ser el pilar fundamental en todo lo que soy, en toda mi educación, tanto académica, como de la vida, por su incondicional apoyo perfectamente mantenido a través del tiempo. Todo este trabajo ha sido posible gracias a ellos.

Finalmente agradezco en especial a mi esposa por sus palabras y confianza, por su amor y brindarme el tiempo necesario para realizarme profesionalmente, a mis amigos, primos, compañeros y a todas aquellas personas que de una u otra manera han contribuido para el logro de mis objetivos.

## CONTENIDO

	pág.
INTRODUCCIÓN.....	9
1 DEFINICIÓN DEL PROBLEMA .....	12
1.1 ANTECEDENTES DEL PROBLEMA.....	12
1.2 PREGUNTA DE INVESTIGACIÓN.....	13
2. OBJETIVOS .....	14
2.1 OBJETIVO GENERAL.....	14
2.2 OBJETIVOS ESPECÍFICOS .....	14
3. JUSTIFICACIÓN .....	15
4. DISEÑO METODOLÓGICO .....	17
4.1 TIPO DE INVESTIGACIÓN .....	17
4.2 DELIMITACIÓN Y ALCANCE .....	18
4.3 POBLACIÓN DESTINATARIA.....	19
5. MARCO TEÓRICO .....	20
5.1 LENGUAJES DOCUMENTALES .....	20
5.1.1 Definición de lenguajes documentales.....	20
5.1.3 Evolución de los lenguajes documentales.....	23
5.1.4 Lenguaje sintético y lenguaje natural controlado. ....	30
5.1.5 Ventajas y desventajas del lenguaje natural controlado y lenguaje sintético.....	33
5.1.6 Sistemas híbridos.....	35
5.1.7 Terminología y término.....	36
5.1.8 Términos de indización.....	37
5.1.9 Términos simples y compuestos. ....	37
5.1.10 Términos en singular y plural.....	39
5.1.11 Accidentes gramaticales: sinónimos y polisémicos.....	41
5.1.12 Forma de presentación de los términos:.....	42
5.1.13 Selección de términos.....	46
5.1.14 Definición de tesoro. ....	53
5.1.15 Historia del tesoro. ....	56

5.1.16	Funciones y uso del tesoro. ....	60
5.1.17	Construcción de tesauros. ....	61
5.1.18	Definición de término descriptor y no descriptor ....	62
5.1.19	Características de los tesauros. ....	65
5.1.20	Relaciones del tesoro. ....	67
5.1.21	Relaciones y aplicación de los símbolos del tesoro. ....	68
6.	CONSTRUCCIÓN DEL TESAURO .....	73
6.1	METODOLOGÍA PARA LA CONSTRUCCIÓN DEL TESAURO EN HIDROCARBUROS .....	73
6.1.1	Profundidad y extensión del Tesoro en Hidrocarburos. ....	73
6.1.2	Fuentes terminológicas. ....	73
6.2	ANÁLISIS E INTERPRETACIÓN DE LA INFORMACIÓN .....	75
6.2.1	Instrumentos de recolección de información. ....	75
6.2.2	Selección y recolección de términos – corpus. ....	75
6.2.3	Establecimiento de facetas y subfacetas. ....	76
6.2.4	Validación de términos. ....	78
6.2.5	Criterios para la construcción de tesauros. ....	80
6.2.6	Reglas semánticas del Tesoro. ....	84
6.3	AUTOMATIZACIÓN DEL TESAURO .....	87
6.3.1	Herramienta para la automatización del tesoro. ....	87
6.3.2	Presentación del tesoro. ....	92
6.3.3	Presentación jerárquica. ....	93
6.3.4	Presentación alfabética. ....	94
6.3.5	Actualización y modificaciones. ....	94
7	CONCLUSIONES .....	97
8	RECOMENDACIONES .....	100
	BIBLIOGRAFÍA .....	100

## LISTA DE TABLAS

	<b>pág.</b>
Tabla 1. Usuarios internos del Centro de Documentación e Información	22
Tabla 2. Usuarios externos del Centro de Documentación e Información	23
Tabla 3. Definiciones de lenguajes documentales	21
Tabla 4. Sinonimia de la coordinación en los lenguajes documentales	31
Tabla 5. Definición de los tipos de lenguajes documentales.	32
Tabla 6. Ventajas y desventajas del uso de los lenguajes documentales según su coordinación.	33
Tabla 7. Términos de indización.	37
Tabla 8. Uso de términos en singular y plural según las normas.	39
Tabla 9. Definiciones sinónimos y polisémicos.	41
Tabla 10. Forma de presentación de los términos.	42
Tabla 11. Selección de términos.	46
Tabla 12. Clasificación de los tesauros según su tamaño.	65
Tabla 13. Clasificación de los tesauros según su idioma.	67
Tabla 14. Relaciones jerárquicas .	68
Tabla 15. Nota de alcance.	69
Tabla 16. Relaciones asociativas.	70
Tabla 17. Relaciones de equivalencia.	71
Tabla 18. Relaciones y aplicación de símbolos del tesauro.	72
Tabla 19. Variables de muestreo.	87
Tabla 20. Símbolo, descripción y definición de NTC 1476:1988.	92

## LISTA DE ANEXOS

	<b>pág.</b>
Anexo A. Formato de extracción y actualización de términos para construcción del tesoro en hidrocarburos.	105
Anexo B. Formato para la evaluación de software para la administración de tesauros	106



## **INTRODUCCIÓN**

Una de las herramientas más utilizadas por las unidades de información, son los lenguajes documentales, teniendo presente que una de sus funciones se focaliza en mejorar los procesos de búsqueda y recuperación de información. Desde esta perspectiva, este tipo de instrumentos para la organización y recuperación de la información, se convierten en un componente fundamental para la normalización de las unidades de información especializadas.

Este proyecto busca mejorar la calidad del servicio que prestan los centros de documentación y las bibliotecas relacionadas con el tema de hidrocarburos, ya que la búsqueda y recuperación de información de manera oportuna es un factor decisivo para la toma de decisiones, la competitividad y el posicionamiento. Por lo tanto, iniciar un proceso de normalización conceptual y terminológica garantiza la pertinencia, eficiencia y oportunidad en la información requerida.

Basados en lo anteriormente enunciado, el presente trabajo tiene como objeto la construcción de un tesoro sobre hidrocarburos y busca potencializarlo como una herramienta útil para los diferentes centros de documentación especializados en el tema que presentan dificultades a la hora de la recuperación de información, debido a la ambigüedad del lenguaje empleado en el análisis y clasificación de la información.

Es en este sentido, que la construcción de un tesoro en hidrocarburos, surge de la necesidad de controlar el uso de términos especializados para describir o representar los documentos técnicos producidos por las empresas del sector petrolero o de hidrocarburos, por las academias asociadas con el sector y por los investigadores.

Este tesoro está enfocado a tres aspectos fundamentales:

1. La indización y visibilidad de los documentos técnicos producidos en este sector.
2. La agilización y mediación en la búsqueda y recuperación de información.
3. El intercambio de información entre las empresas, entidades, academias y demás públicos objetivos interesados en este sector.

Para la elaboración del tesoro se utilizó como directriz la norma técnica colombiana NTC 1476:1998 – Guía para el establecimiento y desarrollo de tesauros monolingües de ICONTEC y su referente internacional, la norma ISO 2788:1986 - Documentation - Guidelines for the establishment and development of monolingual thesauri (Directrices para el establecimiento y desarrollo de tesauros monolingües).

Es importante recalcar que este tesoro es la primera iniciativa para normalizar la terminología especializada en hidrocarburos en idioma español, que cuenta con una estructura taxonómica facetada y que el proceso de exploración efectuado en las diferentes fuentes y recursos de información lo evidencian de tal forma, razón por la cual no fue posible efectuar un referente similar ya que no existe procedencia alguna identificada hasta la fecha.

Es importante recalcar que aunque parte de la terminología de esta área del conocimiento se puede llegar a encontrar en otros idiomas como es el caso del inglés, para nuestro caso solo se abarca el tesoro en este fase en idioma español ya que la misma complejidad de carecer de una estructura previa en este idioma, dificulta su estructuración bajo un escenario multilingüe, por consiguiente esta

primera etapa busca aportar esa primera estructura para que una etapa posterior pueda seguir siendo enriquecido con la terminología en otros idiomas.

El tesauo se encuentra elaborado en la aplicación llamada TemaTres<sup>1</sup>, esta aplicación ha sido diseñada para la gestión y administración de tesauros, es una aplicación libre, amigable de fácil uso, interoperable y que responde a los parámetros descritos en la NTC 1476 y su equivalente ISO 2788.

El tesauo desarrollado se encuentra publicado en un servidor de acceso libre, está disponible a partir de internet de forma abierta y cuenta con una estructura facetada compuesta por 1.130 términos descriptores, 730 relaciones entre los términos, 113 términos equivalentes y 12 notas de alcance.

El tipo de investigación de este proyecto pertenece a la Línea matriz de conocimiento y sociedad y a la Sublínea de la facultad: Bibliotecas, información y sociedad del programa Sistemas de Información y Documentación de la Universidad de la Salle.

---

<sup>1</sup> R020 BIBLIOTECOLOGÍA Y CIENCIAS DE LA INFORMACIÓN - TEMATRES. Software libre para la gestión de tesauros, 2008. [termino de búsqueda: terminología en desastres], [en línea – html] Buenos Aires. [consulta: 30 de nov., de 2008]. Disponible en: <<http://www.r020.com.ar/tematres/>>

# **1 DEFINICIÓN DEL PROBLEMA**

## **1.1 ANTECEDENTES DEL PROBLEMA**

El uso indebido en la asignación de términos para la descripción de los documentos, dificulta y disminuye las probabilidades de recuperar documentos relevantes, factor que aumenta la probabilidad de no disponer de manera inmediata de la información, como fuente primordial para la construcción de nuevos documentos, para la generación de conocimiento y el impulso a la investigación en este sector, para la toma de decisiones vitales y como insumo básico para la ejecución de las tareas y actividades cotidianas.

Para el sector de los hidrocarburos, esta información es de vital importancia, ya que es usada por ingenieros civiles, geólogos, químicos, físicos y de petróleos, entre otros. Los cuales la emplean como fuente de consulta y referencia para la elaboración de sus actividades diarias. Esta situación hace necesaria la construcción de este instrumento con el fin de orientar sus procesos de búsqueda y localización de información, la transferencia del conocimiento generado tanto explícito como implícito, la mediación comunicativa ante un sector tan especializado y de alto impacto económico, político y social y por consiguiente la inter y transdisciplinar como motor de cambio, crecimiento y transformación social.

## **1.2 PREGUNTA DE INVESTIGACIÓN**

¿Existe coherencia y pertinencia en la forma como los profesionales del sector hidrocarburos emplean la terminología y la manera como esta se encuentra registrada en los documentos que guían y direccionan el quehacer disciplinario?

## **2. OBJETIVOS**

### **2.1 OBJETIVO GENERAL**

Construir un tesoro especializado sobre hidrocarburos en el contexto Colombiano, con el fin de normalizar la terminología y permitir la recuperación de los temas propios de este campo del conocimiento.

### **2.2 OBJETIVOS ESPECÍFICOS**

- Contribuir a la normalización del lenguaje documental en el campo de los hidrocarburos.
- Facilitar la descripción, la búsqueda y la recuperación de información en el sector hidrocarburos.
- Disponer de un tesoro especializado sobre hidrocarburos en ambiente Web para su consulta libre, como fuente para la normalización terminológica a disposición de todos los interesados.

### **3. JUSTIFICACIÓN**

La inexistencia de control y normalización terminológica que se presenta en la información del sector hidrocarburos, ha causado grandes inconvenientes y obstáculos para la recuperación de la misma, razón por la cual se hace necesario, la construcción de un tesoro especializado y su aplicación, con el fin de lograr la optimización de los procesos de descripción, indización y recuperación de la información y al mismo tiempo establecer control terminológico en ese dominio específico. Permittiéndonos de esta manera, efectuar intercambios de información entre las entidades del sector de hidrocarburos y el desarrollo terminológico para esta área especializada de conocimiento.

Hoy en día es más latente la necesidad de emplear óptimos lenguajes en actividades relacionadas con la organización, clasificación, representación, marcaje y recuperación de información, especialmente en entornos web. Con el fin, de hacerla más asequible, entendible y usada por los usuarios en sus actividades diarias y cotidianas, ya que el fin real de la información se materializa solo en su aplicación y uso y es en este escenario donde realmente adquiere su valor.

Desde la anterior perspectiva, podemos ver implícita nuestra gran responsabilidad social y profesional, de generar potentes herramientas e instrumentos de gestión de información y conocimiento, que nos permitan generar valores agregados a nuestras actividades inherentes al procesamiento de información y que nos apoyen en la construcción de bases de conocimiento social y colectivo, a través del trabajo colaborativo. Es en este escenario donde encuentran gran aplicabilidad

los tesauros, en un ambiente óptimo de democratización de la información y el conocimiento, como lo es la internet.

Es importante recalcar que la información es un bien común que permite a quienes la consultan, obtener oportunidades para mejorar su calidad de vida, pero si este acceso a la información se encuentra precedido de términos que no se encuentran normalizados, difícilmente se logrará democratizar su uso, por esto la correcta disposición y una práctica normalizada para su indización, facilita la recuperación y acceso a información pertinente y relevante como aspecto fundamental para el desarrollo personal, colectivo y social, es aquí en este punto donde se evidencia nuestra función social como sujetos transformadores de las realidades sociales de nuestros entornos y comunidades a partir del suministro de información valiosa que sirva de insumo para la generación de nuevo conocimiento, aportando de esta manera a un mejoramiento de la calidad de vida y al impulso social de nuestras comunidades a partir de individuos adecuadamente informados, bajo principios de equidad e igualdad.



## 4. DISEÑO METODOLÓGICO

### 4.1 TIPO DE INVESTIGACIÓN

Este trabajo se desarrollara desde una perspectiva documental entendida como “estrategia metodológica de obtención de información, que supone por parte del investigador el instruirse acerca de la realidad objeto de estudio a través de documentos de diferente materialidad (escritos, visuales, numéricos, etc.), con el fin de acreditar las justificaciones e interpretaciones que realiza en el análisis y reconstrucción de un fenómeno”<sup>2</sup>.

Es importante recalcar que la construcción de un tesoro, se soporta en gran medida sobre el análisis de la información registrada en los documentos, ya que la extracción terminológica la cual es considerada el insumo fundamental para la construcción de este tipo de instrumentos, solo puede ser captada a partir de un análisis de las fuentes documentales, ya que desde esta perspectiva se hace visible y predecible las palabras empleadas y las posible variaciones acometidas por los diferentes usuarios o productores de la información.

Para la recolección de información se emplearán los fundamentos de la investigación documental como la consulta de fuentes primarias en la materia, la identificación de especialistas en tesoros y en el tema de hidrocarburos; así como las experiencias en investigación cualitativas y cuantitativas.

---

<sup>2</sup> Yuni, José Alberto; Urbano, Claudio Ariel. Técnicas para investigar: recursos metodológicos para la preparación de proyectos de investigación. [en línea]. [Consulta: 20 de Agosto de 2014]. Disponible en: <http://site.ebrary.com/lib/bibliounisallesp/Doc?id=10353071&ppg=102>

Los procesos de comprensión, análisis, interpretación y selección de información relevante y pertinente se especificarán desde el sector de los hidrocarburos y las disciplinas afines.

La búsqueda de la materia prima del tesoro se va a realizar bajo el método deductivo / analítico; este método se basa en el análisis del contenido de un conjunto de documentos identificados, valorados y seleccionados sobre la temática a desarrollar. Este método es un método propio de la construcción de tesauros que suele utilizarse en campos especializados que carecen de referentes normalizados y que permite evidenciar una terminología real en uso.

Según la AFNOR el método deductivo-analítico radica en la agrupación de términos relevantes del lenguaje natural tomados a partir de diversas fuentes, como lo pueden ser:

- Las demandas de los usuarios
- Los propios documentos emitidos desde una especialidad, para este caso hidrocarburos, y
- Los términos indizados por acumulación.

## **4.2 DELIMITACIÓN Y ALCANCE**

Para el desarrollo del presente trabajo de investigación, se tomará como guía la norma NTC 1476 del año 1998 “Guía para el establecimiento y desarrollo de tesauros monolingües”, la cual es una adaptación de la norma internacional ISO 2788 de 1985. El objetivo de esta norma es garantizar una práctica coherente, normalizada y homogeneizada en la construcción de tesauros especializados, aunque las técnicas descritas en esta norma están basadas en los principios

generales aplicables a cualquier campo temático ya que brindan elementos conceptuales aplicables y ajustables a cualquier dominio de conocimiento. Todas ellas razones que sustentan porque se toma esta norma como base para la construcción del tesoro especializado en hidrocarburos.

### **4.3 POBLACIÓN DESTINATARIA**

La población a la cual está orientada este proyecto son las personas que interactúan de una u otra forma con el sector de hidrocarburos, ya sea de forma directa o indirecta, en esta población encontramos dos grandes grupos:

1. Las personas encargadas de los procesos técnicos al interior de las unidades de información y que requieren de herramientas conceptuales como es el caso de los tesauros para emplear y asignar los descriptores más apropiados en la representación de la información contenida en los documentos o recursos de información que reposa al interior de las mismas y que puede ser accesequible a partir de sus bases de datos.
2. Los diversos tipos, categorías y niveles de usuarios que requieren información sobre esta área del conocimiento y por consiguiente, consultan las bases de datos, con el fin de obtener respuesta a sus necesidades, requerimientos y demandas de información, entre los que podríamos mencionar: los investigadores, los ingenieros de las áreas afines, las entidades adscritas a este sector productivo, los estudiantes de los programas afines.

## **5. MARCO TEÓRICO**

### **5.1 LENGUAJES DOCUMENTALES**

Después de consultar la bibliografía especializada en ciencias de la información, se determinó que existen diferentes denominaciones para el desarrollo de instrumentos que ejercen control terminológico en el ámbito de la asignación temática y clasificación de fuentes de información para la representación y recuperación de la información.

Algunos autores utilizan denominaciones como lenguajes de indización, lenguajes controlados, lenguajes normalizados, lenguajes terminológicos, lenguajes de indización y lenguajes documentales, siendo éste último el de mayor uso en la literatura consultada, razón por la cual se ha decidido denominarlo de esta manera.

#### **5.1.1 Definición de lenguajes documentales.**

En la Tabla 3 se presentan distintas definiciones que vislumbran la diversidad de interpretaciones de los lenguajes documentales.

**Tabla 3. Definiciones de lenguajes documentales.**

AUTOR	DEFINICIÓN
GIL URDICIAIN, Blanca:	Todo sistema artificial de signos normalizados que facilitan la representación formalizada del contenido de los documentos para permitir la recuperación manual ó automática de la información solicitada por los usuarios <sup>3</sup> .
RODRÍGUEZ DELGADO, Rafael:	Lenguajes científicos sistematizados y controlados que expresan el contenido de documentos de cualquier clase sirviendo para recuperar la información que se encuentra en ellos <sup>4</sup> .
ESCUELA UNIVERSITARIA DE BIBLIOTECOLOGÍA Y CIENCIAS AFINES:	Sistema de conceptos y/o símbolos destinado a usos bibliotecológicos, estructurado mediante procedimientos de control de vocabulario, con la finalidad de representar mediante esos símbolos ó conceptos el contenido de los documentos, para facilitar su recuperación temática por usuarios indiscriminados <sup>5</sup> .
NTC 1476:1998 – Guía para el establecimiento y desarrollo de tesauros monolingües:	Conjunto controlado de términos extraídos del lenguaje natural y utilizados para representar, de forma breve los temas de los documentos <sup>6</sup> .
NAUMIS PEÑA, Catalina:	Los lenguajes documentales no se utilizarán únicamente para indizar y recuperar, sino que es un elemento organizativo de los

<sup>3</sup> GIL URDICIAIN, Blanca. Manual de lenguajes documentales. Madrid: Ediciones Trea, 2004. P.28

<sup>4</sup> RODRÍGUEZ DELGADO. Rafael. Teoría de sistemas y tesauros facetados. EN : Revista Española de Documentación Científica, vol. 1, No. 3. (dic, 2005) P. 213-224

<sup>5</sup> ESCUELA UNIVERSITARIA DE BIBLIOTECOLOGÍA Y CIENCIAS AFINES. Diccionario de organización y representación de conocimiento, clasificación, indización, terminología. [termino de búsqueda: leguajes documentales]. [en línea-html]. Uruguay. [consulta: 25 ago., 2008]. Disponible en: <[http://www.eubca.edu.uy/diccionario/letra\\_l.htm](http://www.eubca.edu.uy/diccionario/letra_l.htm)>

<sup>6</sup> INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS Y CERTIFICACIÓN. Guía para el establecimiento y desarrollo de tesauros monolingües. Bogotá : ICONTEC, 1998. P.15 (NTC 1476)

Como puede evidenciarse en los autores citados anteriormente, éstos coinciden en que los lenguajes documentales buscan mediante la normalización y el uso de símbolos, la representación del contenido de los documentos para su posterior recuperación de información.

### **5.1.2 Características de los lenguajes documentales**

Los lenguajes documentales deben cumplir con las siguientes características:

- Provee de control terminológico que relaciona estructuras jerárquicas y asociativas en cada uno de los términos dentro de vocabulario especializado.
- Está compuesto por relaciones semánticas mediante la interpretación del significado de un determinado términos.
- Con el uso de términos descriptores para la indización de documentos, se busca representar fielmente el contenido del mismo para la recuperación y uso de la información.
- Deben ser de fácil manejo, y a la vez preservar en lo posible la riqueza informativa del original (homomorfismo).
- Existe mayor precisión en las búsquedas, debido a la normalización terminológica reduce proporcionalmente las ambigüedades terminológicas existentes en un sistema de información.
- Mediante el control terminológico se puede gestionar eficientemente el uso de términos no preferentes para garantizar la recuperación de la información a usuarios no especializados.

---

<sup>7</sup> NAUMIS PEÑA, Catalina. Los tesauros documentales y su aplicación en la información impresa, digital y multimedia. México. Universidad Autónoma de México. 2007. P.113

- Mediante los lenguajes documentales no solo se estructura términos para la indización y recuperación de información, sino también se estructura áreas del conocimiento especializadas.

Los lenguajes documentales desempeñan su propósito de establecer un hilo conductor mediante el uso de términos en lenguaje natural y lenguaje sintético respectivamente, para representar el contenido del documento en la indización y recuperación de la información.

### **5.1.3 Evolución de los lenguajes documentales.**

El hombre es un ser racional, característica que lo diferencia de las demás especies, debido a que éste cuenta con conductas mentales conducentes que facilitan la clasificación y ordenación de la realidad a objetos y elementos abstractos.

“Los seres humanos comunican las representaciones de su conocimiento por vía oral, escrita o gráfica, que se traduce en la utilización de lenguajes, asimismo orales, escritos y gráficos. Fijando la atención, de momento, en los segundos, se distingue entre lenguajes naturales y artificiales”<sup>8</sup>, por ende el conocimiento y la información no son ajenos a su organización mediante las descripciones que facilita su accesibilidad y uso.

La organización de la información no es ajena a la clasificación y ordenación aún más cuando son necesarias para garantizar que la producción intelectual de la humanidad, sea recuperada con el fin de propiciar diálogos, discusiones y generar

---

<sup>8</sup> CURRÁS PUENTES, Emilia. Ontologías, taxonomías y tesauros: manual de construcción y uso. Gijón, Ediciones trea. 2005. P. 26

nuevo conocimiento, siendo éste un proceso cíclico que trae como efecto la evolución intelectual de la sociedad.

La clasificación de la información supone un proceso más complejo, debido a que éste se efectúa mediante características de contenido en los documentos y la ordenación supone ser un proceso físico que consiste en asignar un lugar donde ubicarlo.

Por lo anterior son los lenguajes documentales un instrumento para facilitar la clasificación y ordenación de la información que fue perfeccionándose a través del tiempo. María José López-Huertas Pérez, en su publicación *Lenguajes Documentales: aproximación a la evolución histórica de un concepto*<sup>9</sup>, divide este concepto en tres momentos conceptuales llamados periodos embrionarios, pre científica y científica.

El periodo embrionario inicia en la Edad Antigua hasta el siglo XVI, donde el único desarrollo conceptual que se reconoce son los inicios de las clasificaciones mediante el uso de catálogos de bibliotecas elaborados por los bibliotecarios en respuesta a la necesidad de organizar las colecciones mediante las agrupaciones temáticas.

Según evidencias, una de las primeras clasificaciones bibliográficas data en China en el siglo I a.C. y se le atribuye a un discípulo de Confucio, quién desarrollo un sistema de clasificación basado en divisiones temáticas generales.

---

<sup>9</sup> LÓPEZ-HUERTAS PÉREZ, María José. *EN* : Boletín de ANABAD, Tomo 41, No. 1, 1991. [en línea -PDF]. Madrid, 1991. P 61 – 79. [consultado: 10 de ago., de 2014]. Disponible en: <<http://dialnet.unirioja.es/servlet/extaut?codigo=267457>> ISSN 0210-4164>



Uno de los primeros avances de clasificación bibliográfica en el mundo occidental surgió en siglo III a.C, en la biblioteca de Alejandría, cuando el bibliotecario Calímaco estableció la ordenación por materias y por autores.

En el siglo XVI surge una clasificación metafísica elaborada por Gesner, en la que realiza “una clasificación jerárquica del conocimiento ideada para agrupar las entradas bibliográficas temáticamente”<sup>10</sup>.

Sin lugar a duda, el periodo embrionario constituyó las bases para efectuar la clasificación y ordenación de las fuentes bibliográficas basada en la temática de las mismas, respondiendo de esta manera lógica una necesidad que facilitaba la recuperación de los recursos de información.

En el periodo precientífico, a partir del siglo XVII y comienzos del XIX, surgen iniciativas concretas frente el concepto actual de los lenguajes documentales, debido a que existe una influencia directa de corrientes Europeas de crear un lenguaje que permitiera eliminar las barreras lingüísticas y al mismo tiempo atender la necesidad de encontrar un sistema de clasificación, siendo Seth Ward, Comenius, Kinner, Wilkins y Baillet sus representantes, quienes vislumbraron la importancia de controlar el uso de los sinónimos, homónimos y antónimos; de igual forma establecieron los principios de encabezamientos de materias, desarrollados por Cutter un siglo después.

Para esta época surge la necesidad de establecer control terminológico, lo que constituye el sentido de los lenguajes documentales, como método para minimizar el uso del lenguaje natural por parte de los usuarios, debido a que éste trae

---

<sup>10</sup> Ibid., P. 63

consigo ambigüedades terminológicas que dificultan la recuperación de la información de los recursos bibliográficos.

El periodo científico comprende el siglo XIX hasta el siglo XX, donde surgen las primeras teorías de los lenguajes documentales como respuesta hacia la comprensión de las expresiones lingüísticas de los usuarios mediante el uso de términos preestablecidos.

Los hechos más relevantes frente a los lenguajes documentales es el desarrollado en Estados Unidos a finales del XIX “lo que marcaría el hito en la teoría de acceso temático a la información la Clasificación Decimal de Melvil Dewey y las Reglas para un Catálogo-Diccionario de Charles Ami Cutter. Cada una de ellas se convertirá en el exponente de un tipo diferente de lenguaje documental”<sup>11</sup>, que consiste en un sistema de clasificación decimal, basado en el control terminológico precoordinado y de estructura jerárquica que hasta la fecha sigue vigente.

Para Blanca Gil Urdiciain, “las primeras manifestaciones del lenguaje documental datan de finales del siglo XIX, cuando aparecen las clasificaciones bibliográficas. Estas clasificaciones, inspiradas en la lógica y en los sistemas filosóficos del conocimiento, se fundan en el principio de precoordinación y son de carácter enciclopédico”<sup>12</sup>.

A finales del mismo siglo surge con bastante aceptación principalmente en las bibliotecas de Estados Unidos, la Clasificación de la Biblioteca del Congreso desarrollada y aplicada por el Congreso de esta nación. A la postre se desarrollaron nuevas propuestas de clasificación en diferentes países del mundo,

---

<sup>11</sup> Ibid., P. 66

<sup>12</sup> GIL. Óp. cit., P. 30

el sistema clasificación universal y las listas de encabezamiento de materias se consolidaron porque cruzaron fronteras y se establecieron como un estándar en gran parte de las bibliotecas del mundo, que al día de hoy están presentes como una solución para el control terminológico que prevén por lo general, una única materia asignable a cada documento y resultan bastante útiles para la clasificación y ordenación física de los documentos, sin embargo en la actualidad presenta inconvenientes debido a la constante y diversa producción de información que tiende a especializarse y requiere de mayor abstracción para su recuperación.

Un hecho significativo de este periodo y que marca un referente en los sistemas de clasificación, fue propuesto en 1933 por Shiyali Ramamrita Ranganathan, quien estableció la clasificación por facetas, siendo ésta estructura la base que inspiraría la construcción de tesauros facetados.

Ranganathan consideraba que debía existir una forma de clasificar que fuera más intuitiva para los usuarios, ya que creía que el Sistema de Clasificación Decimal Dewey - SCDD<sup>13</sup> y la clasificación de la Biblioteca del Congreso - LCC<sup>14</sup> usaban términos para indizar que eran muy rígidos y de lenguaje no combinatorio, por esta razón las estructuras facetas propuestas, precisaba una representación del conocimiento más acorde con el contexto y/o dominio sin perder interdisciplinaridad y transdisciplinaridad de la información, como María Cruz afirma que “la clasificación facetada se aplica en un principio a un esquema de

---

<sup>13</sup> MONTGOMERY COUNTRY MEMORIAL LIBRARY SYSTEM. Sistema de Clasificación Decimal Dewey. [termino de búsqueda: SCDD], [en línea - html]. [consulta: 25 de ago., de 2014]. Disponible en :<[www.countylibrary.org/span3.htm](http://www.countylibrary.org/span3.htm)>

<sup>14</sup> THE LIBRARY OF CONGRESS. Library of Congress Classification. [termino de búsqueda: LCC], [en línea - html]. [consulta: 25 de ago., de 2014]. Disponible en :<<http://www.loc.gov/catdir/cpsol/lcc.html>>

clasificación universal y se utiliza posteriormente como método de clasificación en áreas parciales de conocimiento”<sup>15</sup>.

A finales del siglo XX e inicios del XXI periodo científico, la especificidad y el desarrollo del conocimiento ha llegado al punto en que los Sistemas Universales de Clasificación resultan impertinentes como el único lenguaje documental, que “parten de una división del saber hoy superada por la especificidad actual y los dominios particulares de las disciplinas científicas.”<sup>16</sup>.

El lenguaje no combinatorio impide la participación temática de diferentes dominios específicos y novedosos, razón por la cual entorpece la posibilidad de disgregar la información contenida en los documentos y por el contrario segmenta rigurosamente la información a un dominio general, caso similar ocurre con el uso de la Lista Encabezamiento de Materias -LEMB<sup>17</sup>, que son comúnmente utilizadas en las bibliotecas para el uso de términos descriptores del contenido del documento, pero su principal inconveniente según Cruz es el “exceso de precoordinación de los términos y las normas de presentación en los índices. Dichas normas, excesivamente complejas, dan como resultado una falta de criterios homogéneos a la hora de seleccionar los términos, convirtiendo este

---

<sup>15</sup> LAMARCA LAPUENTE, María Jesús. Hipertexto, el nuevo concepto de documento en la cultura de la imagen. [en línea]. Madrid : Universidad Complutense de Madrid. [consulta: 10 de ago., de 2014]. Disponible en: <[www.hipertexto.info/documentos/xfml.htm](http://www.hipertexto.info/documentos/xfml.htm)>

<sup>16</sup> CRUZ RUBIO. María. Análisis documental: indización y resúmenes en bases de datos especializadas. [en línea]. Madrid : CINDOC, 2000. p. 3. [consulta: 14 de ago., de 2014]. Disponible en: <<http://www.iberius.org/es/AisManager?Action=ViewDoc&Location=getdocs:///DocMapCSDOCS.dPortal/2519>>

<sup>17</sup> LEMB. Lista de Encabezamiento de Materias para Bibliotecas. [termino de búsqueda: LEMB], [en línea - html]. [consulta: 25 de ago., de 2014. Disponible en: <[www.accessmylibrary.com/coms2/summary\\_0286-32034427\\_ITM](http://www.accessmylibrary.com/coms2/summary_0286-32034427_ITM)>

lenguaje en un sistema excesivamente “encorsetado” y rígido que impide las combinaciones en la recuperación”<sup>18</sup>.

Sumado a lo anterior, el uso de términos precoordinados acarrea dificultades al usuario para la recuperación de información, debido a que éste utiliza palabras propias del lenguaje natural y “está sometido a las variaciones gramaticales de un mismo término y a las expresiones sinónimas de un mismo concepto”<sup>19</sup>, que no hacen parte del control terminológico extraídos habitualmente de la Lista de Encabezamientos de Materias y utilizados como términos descriptores del contenido del documento, lo que afecta en la búsqueda y recuperación de información que el usuario considera relevante.

Como se evidencia en los tres periodos (embrionarios, precientífico y científico), el surgimiento de los lenguajes documentales se debe a la necesidad de clasificar, ordenar, indizar, recupera y disponer de la información, bajo un control terminológico que reduzca las ambigüedades terminológicas, partiendo de estructuras y aplicación de un dominio.

Los lenguajes documentales a través de la historia se orientan hacia al control terminológico mediante la selección de términos para la indización y la recuperación de la información, que surgió con la uso sistemas de clasificación general (precoordinación) por la literatura existente que correspondía a temas generales del conocimiento.

En el periodo científico aumenta la preocupación del control terminológico, por esto surgen sistemas de clasificación más flexibles orientados a la construcción de

---

<sup>18</sup> CRUZ RUBIO. Óp. cit., P.2

<sup>19</sup> GIL URDICIÁN, Óp. cit., P. 26

estructuras y mapas del conocimiento “que buscan emular los procesos cognitivos humanos en dominios específicos, mediante la creación de modelos que asignan peso a los conceptos y establecen relaciones causales entre ellos”<sup>20</sup>.

La combinación del uso del lenguaje natural es “representada por un vocabulario ilimitado de palabras y frases utilizadas habitualmente en un campo temático”<sup>21</sup>, y el lenguaje sintético basado en fuentes precoordinadas da como resultado los sistemas híbridos tales como los tesauros, taxonomías y las ontologías que combinan eficientemente la precoordinación y la postcoordinación de un dominio específico.

#### **5.1.4 Lenguaje sintético y lenguaje natural controlado.**

Existen diferentes designaciones en los lenguajes documentales según su coordinación, éstas varían de acuerdo a las denominaciones de los autores que abordan este tema, sin embargo su significado no se altera. En la Tabla 4 se expone la sinonimia existente según la coordinación en los lenguajes documentales.

---

<sup>20</sup> LANCASTER, Frederick Wilfrid. Control del vocabulario en la recuperación de información. 2 ed. Madrid: Universidad de Illinois, 1995. P. 177.

<sup>21</sup> TSADIRAS, Athanasios K. Connitive mapping and certainty neuron fuzzy cognitive maps. Information sciences. Volumen. 101. P 109. Citado por; NAUMIS PEÑA, Catalina. Los tesauros documentales y su aplicación en la información impresa, digital y multimedia. México. Universidad Autónoma de México. 2007. P 113.

**Tabla 4. Sinonimia de la coordinación en los lenguajes documentales.**

DENOMINACIÓN UTILIZADA	SINÓNIMO 1	SINÓNIMO 3	SINÓNIMO 4
SINTÉTICO	Precoordinado	Inductivo	Priori
NATURAL CONTROLADO	Postcoordinado	Deductivo	Posteriori
SISTEMA HÍBRIDO	Mixta		

Fuente: El autor.

Autores como Lancaster, Gil y Naumis coinciden que los lenguajes documentales se dividen en dos grupos según su coordinación, entendiéndose éste como el “momento en el cual se realiza la combinación de términos de indización”<sup>22</sup>.

El primero es el lenguaje sintético o también conocido como precoordinado, son “sistemas fijos de clasificación general de todo el conocimiento existente en una época”<sup>23</sup>. Metodológicamente este lenguaje responde a una actividad más sintética sobre el contenido que analítica porque lo que se busca es la captación global de la temática del documento y asignarle un término para la indización y recuperación de la información.

Los casos típicos del lenguaje sintético son;

- Sistemas de clasificación universal.
- Clasificación decimal Dewey.
- Lista de encabezamiento de materias para bibliotecas.

El segundo grupo es lenguaje natural controlado o postcoordinado, el cual corresponde a un proceso de selección de términos directamente de las fuentes

---

<sup>22</sup> NAUMIS. Óp. cit., P.115

<sup>23</sup> NAUMIS. Óp. cit., P.117

documentales a indizar, para validar su uso y establecer las relaciones semánticas por cada término. Este lenguaje cuenta con la particularidad de que sus términos son altamente especializados y validados concisamente por sus usuarios, este rasgo permite que se mantenga la riqueza terminológica de la fuente tanto para la indización como para la recuperación de la información.

Algunos elementos de coordinación de lenguaje natural son:

- Listado de palabras claves
- Listado de descriptores libres.
- Taxonomía
- Tesauro
- Ontologías

En la siguiente tabla se presenta las definiciones para los tipos de lenguajes documentales:

**Tabla 5. Definición de los tipos de lenguajes documentales.**

TIPO		DEFINICIÓN
<b>LENGUAJE</b>	<b>Sintético</b>	Son aquellos que se componen de un vocabulario predefinido que establece unas relaciones entre los términos y que permite la neutralización de la polisemia y sinonimia, facilitando la recuperación de la información. “Lenguajes documentales en los que los términos que lo componen se coordinan en un proceso previo a su utilización” <sup>24</sup> .
	<b>Natural</b>	“Lenguaje natural no significa otra cosa que el lenguaje

<sup>24</sup> CURRÁS. Óp. cit., P. 75



	<b>controlado</b>	del discurso común” <sup>25</sup> , utilizado por los usuarios para búsqueda y recuperación de información. “lenguajes documentales en los que los términos que los componen se coordinan en un proceso previo a su utilización” <sup>26</sup> .
--	-------------------	---

Fuente: El autor.

### 5.1.5 Ventajas y desventajas del lenguaje natural controlado y lenguaje sintético.

Según la coordinación de términos para la recuperación de información, el lenguaje sintético como natural presenta una serie de características que desvirtúan y aprueban el uso para la indización de fuentes y recuperación de información.

A continuación se presentan las ventajas y desventajas por cada uno.

**Tabla 6. Ventajas y desventajas del uso de los lenguajes documentales según su coordinación.**

LENGUAJE NATURAL CONTROLADO	
VENTAJAS	DESVENTAJAS
El uso del vocabulario puede llegar a ser ilimitado de acuerdo al nivel de especificidad del dominio.	Para un mismo concepto pueden existir diferentes denominaciones, debido a que se aceptan todos los términos y en el control de sinónima y polisemia depende de la rigurosidad del mismo.
Los términos utilizados responden a las	Al existir diferentes denominaciones entre

<sup>25</sup> LANCASTER. Óp. cit., P.177

<sup>26</sup> CURRÁS. Óp. cit., P. 75

necesidades terminológicas reales de los usuarios, y estos no deben realizar búsquedas sobre un lenguaje desconocido e infuncional característicos de la precoordinación.	un mismo concepto, pueden existir la posibilidad de establecer relaciones incorrectas entre los términos.
Los usuarios especializados pueden validar estructuras en las cuales exista control terminológico, lo cual favorece la recuperación de información basada en el lenguaje natural controlado. Existe multidimensionalidad de las relaciones entre términos descriptores, lo que amplía el espectro de un dominio y facilita la recuperación de información.	Existen ambigüedades relacionadas con la homografía de los términos, debido a que un término puede escribirse de la misma manera pero tiene una connotación totalmente diferente.
La evolución de la terminología es más específica de acuerdo del área del conocimiento o dominio.	
LENGUAJE SINTÉTICO	
VENTAJAS	DESVENTAJAS
A través del control terminológico se reduce significativamente la ambigüedad semántica mediante el control de los homógrafos, lo que facilita establecer plenamente la representación de un término dentro de un dominio.	El uso de términos basados en la precoordinación responde a métodos de clasificación tradicionales que no se ajustan a las necesidades terminológicas de dominios hoy altamente especializados.
Mayor uniformidad y constancia en la denominación de conceptos.	Generalmente no responden al lenguaje natural controlado requerido por los usuarios y por tanto dificulta efectuar búsquedas en los términos utilizados por los usuarios.

	No existe multidimensionalidad entre términos debido a su estructura no combinatoria entre temáticas.
--	---

### 5.1.6 Sistemas híbridos.

Una vez expuestas las ventajas y desventajas en el uso del lenguaje sintético y natural controlado en la descripción de recursos bibliográficos, existe una metodología que emplea y fusiona estas dos, la cual es conocida como el sistema híbrido o expertos intermediarios, que parte de las estructuras de conocimiento encontradas en la precoordinación con el objeto de sentar bases en las áreas generales del conocimiento, y se complementa con la creación de estructuras a partir de la identificación de los atributos terminológicos sobre las fuentes de información a las cuales se aplicará la indización y la recuperación.

Con base en los sistemas de clasificación que sentaron los fundamentos en la descripción de información como el Sistema de Clasificación Universal, Clasificación Decimal Dewey y la Lista de Encabezamiento de Materias constituyen en punto de partida para la recolección de términos especializados en un dominio en particular.

Esta labor puede parecer tediosa, sin embargo es pertinente recordar que los lenguajes documentales no se limitan a la indización y recuperación de la información, sino que constituye en un elemento organizativo de contenidos documentales a través de estructuras y mapas de dominios que puede ser generales como especializados.

### 5.1.7 Terminología y término.

#### **Definición.**

Disciplina que lingüística que estudia el comportamiento, cambios, evoluciones y/o adaptaciones que van sufriendo los términos de un dominio que evoluciona naturalmente a medida que aumenta la especialización en el conocimiento científico.

Por lo anterior constituye en la ciencia que se ocupa de la recopilación, el estudio sistemático, la normalización, la interrelación y el enriquecimiento de los vocabularios, términos científicos de lenguajes especializados, donde el término es la unidad mínima del lenguaje escrito o hablado que representa entes concretos y abstractos.

Es aquí donde surge la importancia de los lenguajes documentales, porque analiza los términos lingüística y gramaticalmente, con el fin ubicarlo dentro de un dominio y/o estructura de conocimiento, para el uso efectivo del “término en la indexación y recuperación información en el marco del Bibliotecología”<sup>27</sup>.

Para lograr lo anterior es necesario realizar una serie de actividades, de tal forma que debe garantizarse que el uso del término cumpla con el análisis suficiente para que su uso, responda a la efectiva transferencia del contenido de información.

---

<sup>27</sup> NAUMIS. Óp. cit.,P.170

### 5.1.8 Términos de indización.

Los conceptos susceptibles a representar mediante el uso de términos de indización, se agrupan en tres clases que se exponen a continuación.

**Tabla 7. Términos de indización.**

Términos de indización		
Clase	Alcance	Ejemplo
Entes concretos	Cosas y sus partes físicas	Roca
	Materiales	Cemento
Entes abstractos	Acciones y eventos	Emergencia
	Entidades abstractas y propiedades de cosas, materiales o acciones	Glaciación
	Disciplinas o ciencias	Sismología
	Unidades de medida	Escala sismológica de Richter
Entes individuales	Expresados por nombre propios.	Colombia

### 5.1.9 Términos simples y compuestos.

Los términos se clasifican en simples y los compuestos.

- **Términos Simples:**

Son caracterizados por el uso de un solo término y representan de manera inequívoca un concepto por la condición de su claridad y flexibilidad en la utilización. Por lo anterior no requieren de aclaración

Ejemplo:

Epicentro

- **Términos Compuestos:**

Son términos compuestos por grupos de términos que componen una unidad lingüística que a mayor combinación de elementos terminológicos aumenta la precisión y representación del conocimiento.

Ejemplo:

Flujo del suelo

Canalización de aguas

Amenaza por sismo

Factor detonante natural

Los términos compuestos están integrados por un foco o sustantivo que identifica la clase o categoría y el modificador que establece la característica o diferencia que precisa su significado.

Ejemplo:

**Fenómenos geológicos**

**FENÓMENOS**      Foco

**GEOLÓGICOS**    Modificador

Parar los lenguajes documentales, es recomendable que los términos compuestos mantengan su originalidad morfológica, ya que en caso de dividir los términos pueden afectar su comprensión y puede propiciar pérdida de su significado.

### 5.1.10 Términos en singular y plural.

A continuación se presentan un paralelo de conceptos entre la Lista Encabezamiento de Materias: normatividad para su redacción<sup>28</sup> y la Norma UNE 50-106-90 en lo que respecta al uso del singular y plural:

**Tabla 8. Uso de términos en singular y plural según las normas.**

Clase	ENCABEZAMIENTO DE MATERIAS	UNE 50-106-90	Ejemplo
Singular	Conceptos abstractos.	Conceptos abstractos.	Personalidad.
	Fenómenos de la naturaleza.		Lluvia.
	Propiedades, condiciones o características.	Propiedades.	Temperatura.
	Procesos o actividades ejercidas o sufridas por los seres vivos.	Actividades y disciplinas.	Fotosíntesis.
	Sistemas de creencias o doctrinas.	Sistemas de creencias.	Comunismo.
	Entes concretos no cuantificables.		Sal.
	Partes u órganos del cuerpo singulares o únicos.	Si es órgano único.	Tórax.
Plural	Grupo de seres vivos, entidades u objetos reales.	Cuando se considera un concepto abstracto como representante de una clase con más de un miembro.	Bibliotecas.

<sup>28</sup> ENCABEZAMIENTOS DE MATERIA: normativa para su redacción. Madrid: Biblioteca Nacional, Ministerio de Cultura, 1991. P. 17-20.

	Grupo de personas como grupos étnicos, nacionales, religiosos y nombres de profesionales.		Colombianos.
	Partes y órganos del cuerpo dobles o múltiples.		Ojos.
	Epígrafes de forma.		Diccionarios.
Singular y plural	Cuando las palabras en singular y plural tengan diferente significado.		Derecho civil. Derechos civiles.
	Cuando se quiera diferenciar una técnica de sus productos.		Estadística. Estadísticas.

Ambas normas coinciden en que se debe procurar el uso en un lenguaje documental en singular.



### 5.1.11 Accidentes gramaticales: sinónimos y polisémicos.

En los lenguajes documentales la terminología usada debe ser inequívoca y esto se logra mediante efectivo control de los accidentes gramaticales que se presenta en el uso de sinónimos y polisémicos.

**Tabla 9. Definiciones sinónimos y polisémicos.**

CLASE	DEFINICIÓN	EJEMPLO
Sinónimos	Son palabras que representan un solo concepto y su utilización puede ser favorable cuando este precedida de control terminológico, ya que sirven como puntos de acceso en el sistema de recuperación.	Sismo Terremoto Movimiento telúrico
Polisémicos	Son aquellos que cuentan con "multiplicidad de aceptación de una palabra" <sup>29</sup> o significación múltiple y requieren el complemento para delimitar su contexto.	Rio (Recurso hídrico). Foco (punto de origen de sismo). Foco (punto donde convergen los rayos de luz)).

---

<sup>29</sup> GIL. Oc. cit., P. 191

### **5.1.12 Forma de presentación de los términos:**

Existe un conjunto de reglas semánticas que son necesarias para establecer pautas frente al uso de términos descriptores, por esta razón se presenta una tabla comparativa entre la Norma UNE 50-106-90 y el Libro Manuales de lenguajes documentales de Blanca Gil Urdiciain.

#### **Tabla 10. Forma de presentación de los términos.**

CLASE	UNE 50-106-90	MANUALES DE LENGUAJES DOCUMENTALES
Sustantivos y frases sustantivadas	<p>Los términos de indización deben ser preferentemente sustantivos o frases sustantivadas. Las frases sustantivadas pertenecen a la categoría de términos compuestos y aparecen en dos formas:</p> <p><b>Frases adjetivas</b></p> <p>Esta clase incluye también aquellos términos compuestos por una sola palabra que pueden ser descompuestos morfológicamente en un sustantivo más un modificador con función de adjetivo.</p> <p><i>Ejemplos:</i></p> <p>PÁJAROS MARINOS BALONMANO</p> <p><b>Frases preposicionales</b></p> <p>Las partes de un término compuesto que funcionan como modificadores (véase la parte b en la nota del Apartado 3 .7) se deben considerar como fuentes de términos adicionales en un tesoro. Cuando el modificador contiene un adjetivo, debe preferirse como término candidato el sustantivo del que deriva.</p> <p>Cuando tales términos se acepten, el tesoro debe indicar reenvíos recíprocos entre el término adicional y el término compuesto:</p> <p>PÁJAROS MARINOS TR MARES <b>MARES TR PÁJAROS MARINOS</b></p>	<p>Bajo frases nominales la norma entiende los términos compuestos: frases preposicionales o frases preposicionales</p>

<p>Adjetivos</p>	<p>Debe evitarse, en lo posible, el uso aislado de adjetivos en un lenguaje de indización, aunque tal uso puede darse en circunstancias especiales como las descritas a continuación.</p> <p>Pueden aceptarse adjetivos, como palabras independientes en un índice o tesoro, en las situaciones siguientes:</p> <p>a) cuando en un idioma los adjetivos preceden generalmente a los sustantivos que modifican, se puede hacer un reenvío desde un sustantivo a un adjetivo que figura como primer componente de varios términos compuestos. Por ejemplo en inglés, podría hacerse un reenvío desde el sustantivo "France" hacia el adjetivo "French" si el lenguaje de indización contiene términos como "French art", "French language", "French literature". Esto es especialmente válido cuando la ortografía del sustantivo y el adjetivo varían, por ejemplo : France/French, Sea/Marine ;</p> <p>b) en idiomas en que los adjetivos siguen a los sustantivos que califican puede hacerse un reenvío desde el adjetivo hacia una ó más de las frases nominales que lo contengan. Un índice castellano, por ejemplo, podría contener reenvíos desde "Quirúrgico" hacia términos compuestos como "Equipo quirúrgico", "Instrumental quirúrgico", "Infección quirúrgica", etc.</p>	<p>El empleo de adjetivos aislados debe ser evitado. Los adjetivos de frases adjetivas pueden dar lugar a reenvíos.</p> <p>Ejemplo: FANERÓGAMAS (véanse las diferentes plantas fanerógamas)</p>
<p>Adverbio</p>	<p>En otros idiomas como el inglés, adverbios como "very" ó "highly" pueden no ser usados únicamente como términos indexados. Una frase que comienza con un adverbio puede no ser receptada como un término de indización</p>	<p>Adverbios tales como <i>muy</i> o <i>grandemente</i> no son admitidos, salvo si la forman parte de expresiones</p>

	excepto cuando esta adquiriera un significado especial dentro de una jerga.	técnicas. Ejemplo: MUY ALTA FRECUENCIA
Verbos	Los verbos expresados como participativos o en infinitivos no deben utilizarse como términos de indización. Las actividades deben presentarse como sustantivos o formas verbales sustantivadas. Ejemplo: ADMINISTRACIÓN (no "administrar " )	El infinitivo de los verbos no debe ser empleados solo como término de indización. Las actividades deben ser sustantivadas. Ejemplo: DECORADO (no decorar)
Abreviaturas, siglas y acrónimos	No se debe utilizar abreviaturas, siglas y acrónimos como términos preferidos excepto cuando son ampliamente usados y se reconocen fácilmente dentro del campo cubierto Muchos acrónimos y abreviaturas se pueden referir a más de un concepto y la forma completa del nombre debe funcionar como el término preferente, con una referencia recíproca desde la forma abreviada. Pueden admitirse las siglas y acrónimos como términos preferentes cuando su uso se ha consagrado tanto que resulta rara o se desconoce la forma completa del nombre. Aun así, deben mantenerse los reenvíos recíprocos entre una y otra forma. ORGANIZACIÓN MUNDIAL DE LA SALUD UP OMS <b>OMS USE ORGANIZACIÓN MUNDIAL DE LA SALUD</b>	Es preferible utilizar la forma desarrollada. Las siglas y acrónimos deben evitarse, salvo si se han sustantivado (LÁSER, RADAR) o si se conocen en cualquier idioma (UNESCO)

### **5.1.13 Selección de términos.**

Para que exista un criterio de selección de términos, es preciso conocer las posturas expuestas por la Norma UNE 50-106-90 y la Guía Metodológica para la Elaboración de tesauros creada por el ICFES.

#### **Tabla 11. Selección de términos.**

CLASE	UNE 50-105 – 90	GUÍA METODOLÓGICA PARA LA ELABORACIÓN DE TESAUROS.
Ortografía.	Debe adoptarse la ortografía más generalmente aceptada de cada palabra. Si existen y se aceptan variaciones ortográficas de un mismo término, todas deben introducirse en el tesoro, con reenvíos desde las formas no preferentes a la preferente.	
Extranjerismo y barbarismo	A veces se encuentran términos tomados de otras lenguas, es decir, términos extranjeros aceptados como de nuevo cuño. Si tales términos están bien establecidos, deben incorporarse al tesoro. En ocasiones coexisten un extranjerismo y su traducción. Debe considerarse preferente el extranjerismo si su aceptación es más general; pero si la traducción está bien establecida, debe considerarse preferente. Deben indicarse reenvíos recíprocos entre los dos términos, preferente y no preferente.	A menudo es útil dar los términos equivalentes en otras lenguas para asegurar el uso correcto del descriptor en el análisis de textos en lengua extranjera. Las traducciones deben tratarse como sinónimos o cuasi-sinónimos.
Transliteración (Castellanización)	Cuando se incorporan términos de idiomas con alfabetos no latinos, debe seguirse el	Se hace referencias cruzadas (UP), seleccionando la palabra con la ortografía

	procedimiento descrito en las normas pertinentes.	castellana y remitiendo a ésta el término de la ortográfica extranjera.
Términos de jergas y argot	<p>Cuando existe un concepto de reciente aparición originado en un grupo o subcultura particular y no se halla un término alternativo de aceptación general, el término de argot debe aceptarse como término de indización.</p> <p>Ejemplo:          BYTE          HIPPIES</p> <p>b) Si un término de argot aparece como alternativa a un término ya existente, bien establecido, el término existente debe considerarse preferente y el de argot debe admitirse sólo en el caso de que su uso esté extendido y pueda servir como punto de acceso al índice para el usuario.</p> <p>Ejemplo:          NEOPLASIAS UP neos          neos USE NEOPLASIAS</p>	Cuando un concepto es expresado mediante un término originado dentro de una subcultura, grupo social o región en particular y no existen otra alternativa comúnmente aceptada, este debe ser aceptado, como término de indización.
Denominaciones	Con frecuencia, un producto se conoce por su	Frecuentemente a un producto se le conoce a



comunes y comerciales	<p>nombre comercial</p> <p>En el caso de que exista también un nombre común apropiado, éste se debe adoptar como término preferente y el nombre comercial se aceptará sólo si puede servir como punto de acceso al índice para el usuario.</p> <p>Ejemplo:          ÁCIDO ACETILSALICÍLICO UP aspirina          aspirina USE ÁCIDO ACETILSALICÍLICO</p>	<p>través de una marca de fábrica ampliamente conocida. Cuando también existe un nombre comúnmente apropiado, este último debe ser adoptado como término preferido (descriptor) y la marca de fábrica aceptada como término no preferido (no descriptor), solo si puede ser utilizado por el usuario como punto de acceso para la recuperación.</p>
Denominaciones coloquiales y nombres científicos	<p>Denominaciones coloquiales y científicas.</p> <p>Cuando un mismo concepto se designa por un término coloquial y un término científico se debe elegir la forma más próxima al pensamiento de los usuarios. Por ejemplo, "Pingüinos" puede ser considerado preferente en un índice general, mientras que su equivalente científico, "Esfenisciformes", puede ser preferente en un índice zoológico. En estos casos deben hacerse reenvíos recíprocos.</p>	<p>Cuando un nombre popular y uno científico se refiere al mismo concepto, se debe preferir aquella forma por la cual el usuario realice la búsqueda de información. Por ejemplo: pingüinos puede ser elegido como término preferido (descriptor) en un índice general, pero su equivalente científico "Esfenisciformes" puede preferirse en una consulta o búsqueda en el área de la zoología. En este caso, se debe elaborar referencias cruzadas, entre los términos de acuerdo con el tipo de relaciones establecidas.</p>

<p>Nombres de lugares (Toponímicos)</p>	<p>Los nombres de países y regiones geográficas varían en ocasiones entre idiomas. También en una comunidad monolingüe pueden encontrarse variantes relativas al mismo lugar, por razones como las siguientes:</p> <p>a) Uso común tanto de la denominación "oficial" como de la "popular"</p> <p>Ejemplo: HOLANDA (PAÍSES BAJOS)</p> <p>b) Coexistencia de las formas original y vernácula</p> <p>Ejemplo: PAMPLONA (IRUÑA)</p>	<p>Algunas veces los nombres de los países y regiones varían de idioma a idioma. Así mismo dentro de una comunidad de lenguas aparecen términos diferentes, que se refieren a un mismo lugar. Como regla general, se debe considerar como término preferido (descriptor) aquel que sea de mayor uso entre los usuarios del tesoro. Se deben establecer referencias entre las formas preferidas y no preferidas.</p>
<p>Nombres propios de instituciones y personas</p>	<p>Con frecuencia, estos nombres se excluyen de un tesoro.</p> <p>Si se incluyen, se debe seleccionar la forma del nombre de acuerdo con los principios de registro y codificación prescritos. A veces los nombres pueden variar según su origen ; en estos casos se deben seguir los siguientes principios:</p>	<p>Generalmente existe tipo de términos que se excluyen de un tesoro. Cuando se decide incluir, la forma del nombre debe seleccionarse con base a los principios establecidos en las reglas angloamericanas de catalogación.</p>

	<p>a) Los nombres de instituciones locales o nacionales, es decir, aquellas que llevan a cabo sus actividades y publican quizás revistas o actas en un lenguaje nada más, se deben registrar en su idioma original. Si existe una denominación traducida, debe considerarse no preferente.</p> <p><i>Ejemplo:</i></p> <p>AMERICAN MEDICAL ASSOCIATION  UP Asociación Médica Americana  Asociación Médica Americana  USE AMERICAN MEDICAL ASSOCIATION</p> <p>b) Los nombres de organismos internacionales o de instituciones locales que publican documentos en más de un idioma deben expresarse en su forma más conocida. Si el índice es empleado por usuarios de diferentes comunidades lingüísticas, que pueden conocer el nombre "local" de una organización internacional, pueden admitirse, como no preferentes, las denominaciones vernáculas.</p> <p><i>Ejemplo (en un tesauro castellano):</i></p>	
--	--	--

	<p>ORGANIZACIÓN DE LAS NACIONES UNIDAS  UP UNITED NATIONS ORGANIZATION  UP ONU  United Nations Organization  USE ORGANIZACIÓN DE LAS NACIONES UNIDAS  ONU  USE ORGANIZACIÓN DE LAS NACIONES UNIDAS</p> <p>c) Los nombres propios se registran usualmente en su forma original. Algunos individuos, en especial figuras históricas que han alcanzado renombre internacional, tienen nombres "locales" tan bien establecidos que deben aceptarse como preferentes, estableciendo los reenvíos entre los nombres originales. Debe tenerse en cuenta, sin embargo, que esto resulta excepcional.</p> <p><i>Ejemplo (en un tesauro castellano):</i>  JUANA DE ARCO  UP Jeanne d 'Arc  Jeanne d'Arc USE JUANA DE ARCO</p>	
--	---	--

#### 5.1.14 Definición de tesoro.

En la tabla que se presenta a continuación se presentan algunas definiciones del término tesoro.

**Tabla 12. Definiciones del término “Tesoro”.**

AUTOR	DEFINICIÓN
ISO 2788 Directrices para el establecimiento y desarrollo de tesoros monolingües:	“Según su función: instrumento de control terminológico que traduce a un lenguaje sistémico o documental el lenguaje natural empelado en los documentos y por los usuarios. Según su estructura: un vocabulario controlado y dinámico de términos relacionados semántica y jerárquicamente, que se aplica a un campo específico del conocimiento” <sup>30</sup> .
ESCUELA UNIVERSITARIA DE BIBLIOTECOLOGÍA Y CIENCIAS AFINES:	“Tipo de lenguaje documental que se integra con términos analizados y normalizados que guardan entre sí relaciones semánticas y funcionales. El tesoro se organiza bajo fuerte control terminológico, con objeto de proporcionar un instrumento idóneo para el almacenamiento y la recuperación de la información en áreas especializadas. Puede ser monolingüe, monolingüe con equivalencias o multilingüe, conforme a la cobertura idiomática que proponga. En ciertos casos, agrega una notación. // 2. Repertorio que inventaría, con la aspiración de exhaustividad, el conjunto de unidades léxicas de una lengua. Se diferencia del diccionario general de una lengua en que recopila diacrónicamente y sin criterios de selectividad” <sup>31</sup> .

<sup>30</sup> ORGANIZACIÓN INTERNACIONAL DE NORMALIZACIÓN. Documentation - Guidelines for the establishment and development of monolingual thesauri: ISO, 1986, P. 30. (ISO 2788)

<sup>31</sup> [http://www.eubca.edu.uy/diccionario/letra\\_t.htm](http://www.eubca.edu.uy/diccionario/letra_t.htm), Op., cit.

SLYPE, Georges Van:	“Es una lista estructurada de conceptos, destinados a representar de manera unívoca el contenido de los documentos y de las consultas dentro de un sistema documental determinado, ya ayudar al usuario en la indización y en la correcta recuperación de la información consultada” <sup>32</sup> .
UNESCO - UNISIST	Según su función: son un instrumento de control terminológico, usado para trasladar, desde un lenguaje natural de los documentos, los descriptores, a un sistema lingüístico. Según su estructura: son vocabularios controlados y dinámicos de términos relacionados semántica y genéricamente que cubren un dominio específico de conocimiento.
CORDÓN, J.: LÓPEZ, Lucas.: VAQUERO, J. manual de investigación bibliográfica y documental. Teoría y práctica. Madrid: Ediciones Pirámide, 2001.	Desde el punto de vista de estructura, es un vocabulario controlado y dinámico de términos que tiene entre sí relaciones semánticas y genéricas y que se aplica a un campo particular del conocimiento. Desde el punto de vista funcional, es un instrumento de control de la terminología, utilizado para trasladar, en un lenguaje más estricto, el lenguaje utilizado en los documentos.
GIL URDICIAIN, Blanca:	Lenguaje documental de estructura combinatoria, de carácter especializado, que se basa en expresiones conceptuales llamadas descriptores, provistas de relaciones semánticas.
NAUMIS PEÑA, Catalina:	Una representación de la estructura conceptual de un campo temático; la aclaración del sentido único en el cual se debe utilizar el término (para un sistema de información específico); “la eliminación de los equívocos mediante el envío a términos claros de equivalencia univalente; y las referencias recíprocas entre términos relacionados” <sup>33</sup> .

<sup>32</sup> SLYPE, Georges Van. Los lenguajes de indización: concepción, construcción y utilización de los sistemas documentales. Madrid : Fundación Germán Sánchez Ruipérez, 1991. P. 21.

<sup>33</sup> NAUMIS. Op. cit., P.125

<p>INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS Y CERTIFICACIÓN</p>	<p>Desde el punto de vista de su función, un tesoro es una herramienta de control terminológico utilizado para llevar el lenguaje natural a un lenguaje documental (lenguaje de información) y viceversa.</p> <p>Desde el punto de vista de su estructura, el tesoro es un vocabulario controlado y dinámico de términos que guardan entre sí relaciones semánticas, el cual se aplica en forma exhaustiva a un campo específico del conocimiento<sup>34</sup>.</p>
--	---

Fuente: El autor.

Reuniendo las propiedades anteriormente mencionadas por los diferentes autores, se puede definir la palabra tesoro como: según su función es el resultado de una actividad intelectual mediante la cual se establece el uso de los lenguajes documentales a partir del análisis del lenguaje natural donde el usuario encuentra deducciones semánticas que le permiten la recuperación de información y según su estructura es un control terminológico en el que se establecen relaciones semánticas tanto en niveles como categorías en un dominio específico.

---

<sup>34</sup> INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS Y CERTIFICACIÓN. Óp. cit., P.66

### 5.1.15 Historia del tesoro.

El término tesoro proviene del latín Thesaurus ó Tesoro, sin embargo “existe gran variedad de opiniones sobre los antecedentes etimológicos de la palabra como conceptuales del tesoro. La mayoría de los autores se limitan a dar cuenta de su raíz semántica griega como tesoro, bien material ó del conocimiento”<sup>35</sup>.

El término fue usado inicialmente para referirse a diccionarios e índices “con la misma connotación de lugar donde se guarda lo más importante”<sup>36</sup>. El primero que uso la palabra “tesoro” en el campo de la literatura fue “Florentino Brunetto Latini en su enciclopedia *Livre du Trésor*”<sup>37</sup>, en el siglo XIII. En la lengua española el primer diccionario fue creado por “Sebastián de Covarrubias reimpresso en 1672 por Gabriel de León”<sup>38</sup>, sin embargo quien popularizo el término “tesoro” fue Peter Mark Roget nacido en 1779 quien retirado de su vida profesional empieza esta ardua labor que se mantiene en la memoria con su obra “*Thesaurus of english words and plirases classijed and arranged so as lo facilitate the expression of ideas and assist itt literary composition*”<sup>39</sup>, publicado en 1852.

Roget intentaba presentar la distinción existente entre un diccionario de un tesoro, debido a que el primero cuenta con un orden alfabético de palabras sin

---

<sup>35</sup> RUBIO LINIERS, Maria Cruz. Tesoro de historia contemporánea de España. Madrid: Consejo superior de investigaciones científicas. 1999. 57 p. Citado por; NAUMIS PEÑA, Catalina. Los tesoros documentales y su aplicación en la información impresa, digital y multimedia. México. Universidad Autónoma de México. 2007. P 122.

<sup>36</sup> *Ibíd.*, P.122

<sup>37</sup> *Ibíd.*, P.123

<sup>38</sup> *Ibíd.*

<sup>39</sup> ROGET, Peter Mark. Thesaurus of English words and Plirases Classijed and Arranged so as lo Facilitate the Expression of Ideas and Assist itt Literary Composition. [termino de búsqueda: Roget]. [en línea-html]. [consulta: 12 de nov., de 2014]. Disponible en: <<http://www.bartleby.com/110/>>



ninguna relación y el segundo expresa relaciones entre los términos creando estructuras de conocimiento “que forman categorías conceptuales, presentando las palabras de acuerdo con las ideas que expresan... la palabras del lenguaje están aquí clasificadas, no de acuerdo con su sonido u ortografía, sino estrictamente de acuerdo con su significación”<sup>40</sup>, de esta forma Roget visualizaba tal y como menciona en el título de su obra que al traducirlo indica “*facilitar la expresión de ideas y servir de ayuda en la composición literaria*”, lo cual era el inicio del establecimiento de las relaciones semánticas contenidas en los tesauros.

Precisamente casi un siglo después de la aparición del término “tesauro” en 1957, se usa por primera vez como herramienta para la indización y recuperación de información en un trabajo elaborado por “Helen Brownson, miembro de la American National Science Foundation”<sup>41</sup>, donde su contribución radicó en el análisis de la problemática de traducir los conceptos y sus relaciones, de tal forma que permitiera establecer un lenguaje con mayor precisión y sin ambigüedades, lo que presenta un avance metodológico en las ciencias de información para facilitar la recuperación de la información.

En 1969 Aitchison publica el *Thesauro Facet: a thesaurus and faceted classification for engineering and related subjects* o *Tesauro facetas*, donde presenta “las ventajas de indizar y clasificar con una correspondencia entre facetas y descriptores”<sup>42</sup>, y Aitchison sustentó que su tesauro “adoptaría una

---

<sup>40</sup> GIL URDICIAIN, Blanca. Orígenes y evolución de los tesauros en España. *EN* : Revista General de Información y Documentación. Vol. 8, No. 1 (1998). P. 63 – 110.

<sup>41</sup> ARANO, Silvia. Los tesauros y las ontologías en la Biblioteconomía y la Documentación. [termino de búsqueda: Tesauro y taxonomías]. [en línea-html]. [consulta: 12 de nov., de 2014]. Disponible en: <<http://www.hipertext.net/web/pag260.htm>>

<sup>42</sup> CHAUMIER, Jacques. Análisis y lenguajes documentales: el tratamiento lingüístico de la información documental. Paris: Mitre. 1986. 12 p. Citado por; NAUMIS PEÑA, Catalina. Los

solución mixta mucho más eficaz al aunar la clasificación por materias con una subdivisión por facetas que comprendía dos entradas: la relaciones alfabéticas de los descriptores del tesoro, con las relaciones TE (término específico), TG (término genérico), TR (término relacionado), y el reenvío por un código de tres caracteres a la parte facetada, y la clasificación por materias, en que los términos son reagrupados en facetas o funciones fundamentales<sup>43</sup>, lo que podría sustentarse como localización del término en un contexto alfabético y sistémico o estructural de un dominio.

Entre las década de los setenta y ochenta con la parcelación de la información y la expansión de la terminología cada vez más especializada, la informática empezó a contribuir progresivamente mediante la creación taxonomías electrónicas de conocimiento como recurso para el procesamiento de información de manera estructurada, de tal forma que se dio un paso entre el tradicional tesoro impreso, que por sus características presentaba dificultades para la actualización a un tesoros administrador por computadores que respondían a “modelización matemática”<sup>44</sup>, aplicada a la lingüística, inteligencia artificial y las técnicas de programación, para la indización y recuperación de la información.

Con el auge de construcción de tesoros surgió la necesidad de normalizar y establecer parámetros universales, por esta razón en 1986 la Organización Internacional para la Estandarización (ISO por sus siglas en inglés) publica la norma 2788:1986 - Guidelines for the establishment and development of

---

tesoros documentales y su aplicación en la información impresa, digital y multimedia. México. Universidad Autónoma de México. 2007. P 124.

<sup>43</sup> LÓPEZ, Miguel Ángel. Evolución histórica y tendencias observables en los tesoros. Departamento de Biblioteconomía y Documentación. Madrid. Universidad Carlos III de Madrid. 2006. P 363.

<sup>44</sup> *Ibíd.*, P.365

monolingual thesauri. Posteriormente en el año de 1990 la Asociación Española de Normalización y Certificación publica la UNE 50-106-90 – Directrices para el establecimiento y desarrollo de tesauros monolingües<sup>45</sup>. En el contexto Colombiano el Instituto Colombiano de Normas Técnicas elabora tomando como referencia las normas ISO y UNE para crear la NTC 1476:1998 – Guía para el establecimiento y desarrollo de tesauros monolingües.

Si se considera las iniciativas gestionadas por Roget, Brownson, Aitchison y los organismo internacionales de normalización, se evidencia que la recuperación de la información siempre ha sido un aspecto que ha inquietado ajenos y expertos de la bibliotecología para los procesos de descripción, indización, clasificación, ordenación y recuperación de la información, situación que en la actualidad se viene presentando con la constante creación de más y nueva de información apoyados en los recursos tecnológicos, lo que ha facilitado la apertura de nuevos horizontes para crear, compartir, publicar, aportar, desarrollar desde la Internet, así mismo es irónico pensar que con la creación de herramientas tecnológicas y protocolos mundiales de intercambio y acceso de información como lenguajes de Hyper-Text Markup Language –HTML<sup>46</sup> y Extensive Markup Language - XML<sup>47</sup> entre otros, aún se presenten necesidades de especialistas que contribuyan a la normalización tanto en fuentes de información tanto análogas como digitales, situación en la cual, los tesauros pueden aportar significativamente hacia normalización terminológica de dominios específicos.

---

<sup>45</sup> ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. UNE Directrices para el establecimiento y desarrollo de tesauros monolingües. Madrid: AENOR, 1990, P.40. (50-106)

<sup>46</sup> Lenguaje de Marcas de Hipertexto para la construcción de páginas Web.

<sup>47</sup> Lenguaje de Marcas extensible), es un metalenguaje extensible de etiquetas desarrollado por el World Wide Web Consortium (W3C).

### **5.1.16 Funciones y uso del tesoro.**

Los tesauros cumplen las siguientes funciones:

- Normalizar el vocabulario de un dominio específico a fin de controlar las posibles entradas de búsqueda y recuperación de información mediante el uso del lenguaje natural utilizado por los usuarios.
- Los tesauros no agrupan todos los términos de un dominio, como ocurre en los diccionarios especializados, sino que se encargan de agrupar aquellos que podrían ser utilizados por los usuarios en las consultas, por lo tanto, está concebido en función de las necesidades de los usuarios.
- El tesoro trae inmerso relaciones verticales (jerarquías y/o niveles) y relaciones horizontales (categorías), que permite establecer telarañas entre términos de manera invisible para crear relaciones que facilitan la navegación entre diversos tópicos que complementan y perfeccionan temáticamente las búsquedas del usuario.
- Es un lenguaje documental que responde a la particularidad, que la designación de términos descriptores, específicos y no descriptores que hacen parte de las relaciones jerárquicas, asociativas y equivalencia son propuestas por el bibliotecólogo y/o documentalista en una estructura especializada, y es validada por usuarios especialistas quienes son el fin último que motiva la construcción del tesoro para la recuperación de información.
- Los tesauros además de controlar la terminología existente en un dominio, cumple funciones recurso para la indización, clasificación temática, ordenación de fuentes especializadas y recuperación de información, previa validación del

documentalista y especialista del dominio respectivamente, lo que asegura la calidad en el uso de términos inequívocos como descriptores de temáticas especializadas.

- Es un lenguaje documental que facilita la indización y por ende la clasificación de las fuentes susceptibles de descripción, mediante la representación del contenido de los documentos con el uso de términos descriptores, que cuentan con estructuras invisibles de conocimiento especializado.
- El tesoro debe controlar las posibles entradas utilizada por los usuarios en un sistema de información, de tal forma que brinde alternativas terminológicas que conduzcan a la obtención de resultados satisfactorios que se logran con la eficiente normalización y conformación de estructura y/o taxonomía del dominio especializado.
- Generalmente el usuario desconoce la estructura mental que lo conduce a la temática propia de su necesidad de información, el tesoro mediante sus relaciones semánticas inducen a éste a través de su taxonomía para seleccionar los términos descriptores a ser utilizados en el sistema de información que le permita la recuperación de fuentes que motiva la consulta.

#### **5.1.17 Construcción de tesauros.**

Normas para la construcción de tesauros:

A continuación se presentan las normas para la construcción de tesauros:

- La Organización Internacional para la Estandarización creó la norma ISO 2788 – 1986. E) Documentation - Guidelines for the establishment and development of monolingual thesauri.

- Para la construcción de tesauros multilingües La Organización Internacional para la Estandarización creó la norma ISO 5964 – 1985. Guidelines for the establishment and development of multilingual thesauri.
- En el contexto hispanoamericano existe la norma española UNE 50-106-90 – Directrices para el establecimiento y desarrollo de tesauros monolingües.
- Para el caso Colombiano existe la Norma Técnica NTC 1476:1998 – Guía para el establecimiento y desarrollo de tesauros monolingües.

En los dos últimos casos, las normas hacen aclaración que son equivalentes a la ISO 2788 – 1986 y sus distinciones entre conceptos y metodología son inexistentes, aunque presenta algunas variaciones en el uso de ejemplos y símbolos.

#### **5.1.18 Definición de término descriptor y no descriptor**

- **Término descriptor**

Es un término controlado y normalizado usado para indizar contenidos de documento, labor que facilita la búsqueda y recuperación de la información.

Cada término descriptor hace parte integral de un sistema conceptual que refleja la ordenación y la estructura según el campo de conocimiento en categorías y subcategorías o género - especie.

El término descriptor es “la unidad lingüística cuya función primordial es la descripción de un concepto en el contexto de un ámbito restringido de conocimiento. El conjunto de unidades lingüísticas llamadas descriptores, y sus correspondientes relaciones, conforman el tesauro”<sup>48</sup>.

Catalina Naumis presenta varias consideración asertivas en referencia a las cualidades del término descriptor, las cuales se mencionan a continuación: El contenido o representación del conocimiento (semántica); la definición del objeto que el descriptor designa (terminología y lexicografía); la estructuración dentro de una organización expresada a través de las relaciones conceptuales con otros descriptores (lingüística); la estructura morfosintáctica del descriptor (lingüística); la función para almacenar en un sistema de información (bibliotecología); y la función de recuperador de información para satisfacer una necesidad del usuario (bibliotecología).

La norma UNE 50-106-90, indica que término preferente (descriptor) es “aquel que se utiliza sistémicamente en la indización para representar un concepto determinado, también llamado descriptor”<sup>49</sup>.

Según el ICFES en su Guía para el establecimiento y desarrollo de tesauros monolingües, indica que los descriptores deben responder a las siguientes características:

- Un descriptor debe representar un único concepto. Si la descripción no es suficientemente explícita, se deben usar notas de alcance para evitar ambigüedad.

---

<sup>48</sup> NAUMI. Óp. cit., P.134

<sup>49</sup> ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. Óp. cit., P. 5

- Cuando un descriptor puede ser expresado por varios sinónimos debe escogerse un único término como descriptor, que será aquel que desde el punto de vista del idioma en que se está elaborando el tesoro responda a las siguientes características: “propiedad, concisión y uso”<sup>50</sup>.

#### **a. Presentación de términos descriptores**

Según las normas ISO 2788 de 1986, la UNE 50-106-90 y la NTC 1476: 1998 en su epígrafe 4.4. Literal A. indica.: “los descriptores deben expresarse en mayúsculas”.

- **Términos no descriptores**

Los términos no - descriptores deben incluirse para agregarle valor al tesoro, debido que solo mediante la participación de términos descriptores y no descriptores se logra abarcar las posibles entradas utilizadas por el usuario en la recuperación de la información.

La norma 50-106-90, se refiere a los términos no preferentes (no descriptor) como “sinónimo o cuasi sinónimo de un término preferente (descriptor). Un término no preferente no se asigna a los documentos, pero proporciona un punto de acceso a partir el usuario es dirigido mediante una instrucción (por ejemplo: USE, VÉASE) al término apropiado; el término no preferente se conoce también como término no descriptor”<sup>51</sup>.

---

<sup>50</sup> INSTITUTO COLOMBIANO PARA EL FOMENTO DE LA EDUCACIÓN SUPERIOR. Oc. cit., P.37 – 38

<sup>51</sup> ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. Óp. cit., P.5



**a. Presentación de términos no descriptores:**

Según las normas ISO 2788 de 1986, la UNE 50-106-90 y la NTC 1476: 1998 en su epígrafe 4.4. Literal B.: “los no descriptores deben imprimirse (presentarse) en minúscula, salvo que en los casos de nombres propios que requieran una inicial en mayúscula o las abreviaturas o acrónimos, que siempre deban escribirse en mayúsculas”<sup>52</sup>.

**5.1.19 Características de los tesauros.**

Los tesauros responden a particularidades que lo hacen diferenciados y especializados teniendo en cuenta unos tópicos según:

- 1. Según su tamaño.** Los tesauros crecen de acuerdo al nivel de profundidad ó desarrollo, y la cantidad de términos influyen para que se divida en categorías como se presenta en la siguiente tabla.

**Tabla 12. Clasificación de tesauros según su tamaño.**

CATEGORÍA	DEFINICIÓN	EJEMPLO
MACROTESAURO	“Crear una especie de superestructuras genérica de términos que incluiría a un grupo de tesauros, u otros tipos de vocabularios de diversos campos temáticos” <sup>53</sup> . “Macrotesauro es aquel que contiene diferentes temas	<a href="http://www.uibk.ac.at/info/oecd-macroth/es/">Macrotesauro de la OCDE</a> URL: <a href="http://www.uibk.ac.at/info/oecd-macroth/es/">http://www.uibk.ac.at/info/oecd-macroth/es/</a>

<sup>52</sup> INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS Y CERTIFICACIÓN. Óp. cit., P.5

<sup>53</sup> LANCASTER. Op. cit., P. 215

	relacionados entre sí, pero dispuestos en apartados diferentes, como si se tratase de un conglomerado de tesauros más pequeños <sup>54</sup> .	
TESAURO	Dentro de la definición recolectada en las diferentes fuentes bibliográficas, no indican cuando puede considerarse tesauro según la cantidad de términos que lo componen.	<a href="#">Tesauro de redes de ordenadores</a> URL: <a href="http://www.um.es/gtiweb/fimm/tesauro/">http://www.um.es/gtiweb/fimm/tesauro/</a>
MICROTESAURO	“Tesauro no tiene no tiene definición universalmente aceptada. Con frecuencia se utiliza para hacer referencia a un tesauro pequeño y especializado. Puede definirse como un vocabulario especializado, que es compatible con algún tesauro más amplio y que está totalmente integrado en su estructura jerárquica” <sup>55</sup> .	<a href="#">Tesauro de Huridocs</a> URL: <a href="http://www.huridocs.org/herramientas/HURIDOCsindextermsspahhttp://www.r020.com.ar/eurovoc/index.php?tema=30450">http://www.huridocs.org/herramientas/HURIDOCsindextermsspahhttp://www.r020.com.ar/eurovoc/index.php?tema=30450</a>

Fuente: El autor.

**2. Según su idioma.** Los tesauros según su idioma se clasifican en dos categorías como se presenta a continuación.

<sup>54</sup> CURRAS. Óp. cit., P. 35

<sup>55</sup> LANCASTER. Op. cit., P. 212

**Tabla 13. Clasificación de tesauros según su idioma.**

CATEGORÍA	DEFINICIÓN	EJEMPLO
Monolingües	Los tesauros monolingües están escritos en una sola lengua y están regidos por la norma ISO 2788: 1986. Estos pueden contener términos de extranjerismo que no pueden trasladarse al idioma propio del tesauro porque pierde el significado del término.	Ejemplo: Tesauro en idioma Español <a href="#">Tesauro ISOC de Psicología</a> URL: <a href="http://www.cindoc.csic.es/webpublic/label46.htm">http://www.cindoc.csic.es/webpublic/label46.htm</a>
Multilingües	Está compuesto por términos de diferentes idiomas y están orientados por la ISO 5964: 1985.	Ejemplo: Inglés, Francés, Alemán, Italiano, Español <a href="#">Astronomy Thesaurus</a> URL: <a href="http://msowww.anu.edu.au/librariy/thesaurus/">http://msowww.anu.edu.au/librariy/thesaurus/</a>

Fuente: El autor.

### 5.1.20 Relaciones del tesauro.

Los tesauros deben presentar claramente las relaciones entre sus términos descriptores, que apoyados en el uso de símbolos conforman una estructura de conocimiento, la cual parte de conceptos generales a niveles específicos según su amplitud y especialidad del dominio estudiado.

Sus relaciones se presentan de dos tipos, la primera de ellas son las relaciones jerarquías (verticales) y la segunda las relaciones horizontales (niveles); ambas crean telarañas invisibles entre términos que orientan la navegación entre tópicos que complementan y perfeccionan temáticamente las búsquedas del usuario.

Los tesauros en su estructura cuentan con facetas las cuales constituyen un grupo de términos que representan una característica de división de campo temático. Las facetas (categorías) principales se desglosan en subfacetas (clases) del dominio especializado siguiendo la estructura arborescente, a su vez cuando las subfacetas disponen de términos subordinados, se convierten en facetas hasta el nivel de amplitud temática del tesoro.

#### 5.1.21 Relaciones y aplicación de los símbolos del tesoro.

En los tesauros se reconocen tres tipos de relaciones entre términos:

- a. **Relaciones jerárquicas:** Constituyen en niveles de superordinación y subordinación, donde el primero corresponde al género y el segundo en la especie. Son representadas por los símbolos (TG) y (TE) en el idioma de Español, (BT) y (NT) en el idioma Inglés respectivamente. Su objetivo es expresar su reciprocidad entre los términos genéricos y específicos como se presenta a continuación:

**Tabla 14. Relaciones jerárquicas.**

Fuente: El autor.	SÍMBOLO	DESCRIPCIÓN	DEFINICIÓN
			NTC 1476: 1998
	TG	Término Genérico	El término que sigue representa un concepto con la connotación más amplia.
TE	Término Especifico	El término que sigue se refiere a un concepto con una connotación más precisa.	

Una especie puede ser género para otra especie, lo anterior está determinado por la cantidad de niveles del tesoro.

**Nota de alcance:** Un elemento vital en los términos descriptores especialmente para el indizador, es el uso de las notas de alcance, debido a que su función es la de ser complemento, aclaración, explicación y delimitación de uso de un término descriptor en un contexto temático del tesoro. “La operación por medio de la cual se establecen el nexo entre la expresión lingüística y el objeto que ella designa”<sup>56</sup>.

Es representada por el símbolo (NA) en el idioma de español y (SN) en el idioma Inglés respectivamente.

**Tabla 15. Nota de alcance.**

Fuente:	SÍMBOLO	DESCRIPCIÓN	DEFINICIÓN
			NTC 1476: 1998
El autor.	NA	Nota de Alcance	Una nota que se adjunta a un término para indicar su significado restringido dentro de un lenguaje de indización.

El uso de la nota de alcance, también se aplica para las siguientes condiciones:

- Definición completa de un término.
- Definición restringida destinada a limitar el sentido del término.
- Inclusión de términos en relación jerárquica con el descriptor.
- Historia del término descriptor o nota histórica.

<sup>56</sup> NAUMIS. Óp. cit., P. 195.

**b. Relaciones asociativas:** Son las relaciones existentes entre términos descriptores en multiniveles a nivel horizontal y vertical. Estas relaciones son importante porque dos términos descriptores pueden tener vínculos dentro de la estructura del conocimiento. Es importante La norma NTC 1476:1998 define como aquella que “cubre las relaciones entre pares de términos que no son miembros de un conjunto de equivalencia, ni que puede ser organizados dentro de una jerarquía donde se subordina de otro. Sin embargo, son mentalmente asociados a tal punto que la conexión entre ellos debe hacerse explícita en el tesoro, ya que este recurso ofrece términos alternativos que podrían ser utilizados para la indización o para la recuperación. Esta relación es recíproca”<sup>57</sup>.

Es representado por el símbolo (TR) en el idioma de español y (RT) en el idioma Ingles.

**Tabla 16. Relaciones asociativas.**

SÍMBOLO	DESCRIPCIÓN	DEFINICIÓN
		NTC 1476: 1998
TR	Término Relacionado	El término que sigue está asociado pero no es un sinónimo, cuasi-sinónimo, término genérico o término específico.

Fuente: El autor.

<sup>57</sup> INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS Y CERTIFICACIÓN. Óp. cit., P.40

**c. Relaciones de equivalencia:** Están representado por los símbolos de (UP) y (USE), los cuales cumplen la función de establecer relaciones terminológicas entre términos descriptores y no descriptores.

El control de los términos equivalentes se debe al común uso de los sinónimos y cuasi-sinónimos dentro del léxico utilizado por el indizador para la descripción de la fuente bibliográfica y/o por el usuario para la recuperación de información.

**Tabla 17. Relaciones de equivalencia.**

SÍMBOLO	DESCRIPCIÓN	DEFINICIÓN
		NTC 1476: 1998
UP	Usado por	(Usado por). El término que sigue es un sinónimo o cuasi-sinónimo no preferido.
USE	Término preferido	El término que sigue es el término preferido cuando se debe elegir entre sinónimos o cuasi-sinónimos

Fuente: El autor.

En la tabla que se presenta a continuación sintetiza las tres formas de presentar las relaciones y aplicación de los símbolos del tesauro.

**Tabla 18. Relaciones y aplicación de los símbolos del tesoro.**

RELACIONES SEMÁNTICAS	TIPO DE RELACIÓN	CONCEPTO	CARACTERÍSTICA	SÍMBOLOS	EJEMPLO
	Jerárquica	Es la presentación de un índice sistémico de términos genéricos estructurados en función de las categorías y/o clases siguiendo una jerarquía natural del dominio estudiado, de tal forma que los términos genéricos se presentan en esquema arborescente con ramificaciones sucesivas siguiendo la taxonomía de la faceta del tesoro.	Estructura arborescente.	TG (Término genérico) TE (Término específico)	Género - Especie
	Asociativa	Corresponde a un índice alfabético que presentan los términos descriptores como no descriptores que hacen parte del tesoro con sus respectivas relaciones jerárquicas y de equivalencia.	Relación multidimensional.	TR (Término relacionado)	
	Equivalencia	Es la ponderación que se efectúa entre los términos extraídos mediante la designación de preferencias en el uso de éstos, a través de la localización de sinónimos eliminando las ambigüedad y garantizando que el términos sean inequívocos para la indización y recuperación de la información.	Desviaciones lingüísticas.	UP (Término no preferido) USE (Use)	Sinónimos Cuasi-sinónimos

Fuente: El autor.



## **6. CONSTRUCCIÓN DEL TESAURO**

### **6.1 METODOLOGÍA PARA LA CONSTRUCCIÓN DEL TESAURO EN HIDROCARBUROS**

A continuación se desarrolla la metodología aplicada para la construcción del tesauo en hidrocarburos, construido con el propósito de ser herramienta para la indización y recuperación de información.

#### **6.1.1 Profundidad y extensión del Tesauo en Hidrocarburos.**

El tesauo ha sido construido a partir de terminología utilizada en este sector, y su profundidad está limitada a la estructura y profundidad de la misma, sin embargo el presente tesauo está elaborado para ser empleado y reutilizado en otras entidades que cumplan con una labor semejante, en el contexto Colombiano.

#### **6.1.2 Fuentes terminológicas.**

En el presente trabajo se explicó con antelación varios aspectos necesarios para el procesamiento terminológico y demás elementos característicos de los tesauos.

Teniendo en cuenta que el tesauo está desarrollado para su aplicación en el sector hidrocarburos, gran porcentaje de los términos recolectados se efectuó directamente sobre los documentos técnicos que generan las entidades del sector, tomando como referente ECOPETROL y Ministerio de Minas y Energía, con el

objeto de garantizar que su uso corresponda eficientemente a las labores de indización y recuperación de la información que se administra, en la práctica real.

La terminología es variante y de constante crecimiento, por lo que no se descarta la posibilidad de enriquecer terminológicamente el tesoro, integrarlo y/o reutilizarlo con otros Centro Documentales que estén interesados en esta temática.

Dentro de las fuentes consultadas no se encontró un tesoro construido con el mismo propósito.

Para la construcción del corpus terminológico del tesoro, se tomó la metodología propuesta por el *Sistema Híbrido*, en el cual se agrupa las propiedades del lenguaje sintético y lenguaje natural controlado eficientemente, debido a que elabora la estructura del tesoro inicialmente en áreas generales del conocimiento propuestas por los sistemas precoordinados tales Clasificación Decimal Dewey y la Lista de Encabezamiento de Materias, y se complementa con la metodología propia del lenguaje natural controlado, el cual consiste en la extracción de términos directamente sobre las fuentes técnicas a los cuales posteriormente se le efectuará la indización y recuperación.

## **6.2 ANÁLISIS E INTERPRETACIÓN DE LA INFORMACIÓN**

### **6.2.1 Instrumentos de recolección de información.**

Para la recolección de términos fue utilizada como referencia la planilla de análisis semánticos establecidos por la norma ANSI/NISO Z39.19 – 2003<sup>58</sup>, la cual fue ajustada de acuerdo a la necesidad del proyecto para la recolección, descripción, establecimiento de relaciones entre términos, y se encuentra dispuesta en el Anexo A. del presente trabajo, identificada como *“Formato de extracción y actualización de términos para el tesoro en hidrocarburos”*.

### **6.2.2 Selección y recolección de términos – corpus.**

Por tratarse de un tesoro que utiliza la metodología propuesta por el Sistema Híbrido, la selección y recolección de términos inicia con la consulta de fuentes terminológicas que se basan en un lenguaje sintético (precoordinados), por lo que fueron utilizadas las siguientes referencias:

- Clasificación Decimal Dewey
- Lista de Encabezamiento de Materias para Bibliotecas.

Producto de estas consultas se establecieron las facetas y subfacetas generales.

---

<sup>58</sup> AMERICAN NATIONAL STANDARD INSTITUTE, NATIONAL INFORMATION STANDARD ORGANIZATION. NORMA: ANSI/NISO Z39.19 - 2003. Guidelines for the construction, format and Management of monolingual thesauri. P. 10.

Para complementar la metodología del Sistema Híbrido se efectuó la recolección de términos basado en el lenguaje natural controlado (postcoordinados) directamente sobre los documentos seleccionados como objeto de análisis.

Los términos extraídos se registraron en el “*Formato de extracción y actualización de términos para el tesauro en hidrocarburos*”, éste proporcionó elementos que permitieron el establecimiento de las relaciones jerárquicas, las relaciones de equivalencia, asociativas y las notas de alcance en los términos descriptores que así lo requirieron. De esta manera se elabora la estructura del tesauro en sus facetas y subfacetas.

Los criterios considerados para la selección de términos descriptores fueron:

- Grado de aparición del término en las fuentes susceptibles a indización.
- Valor informativo y descriptivo de un concepto que responde al sector de hidrocarburos.
- Representatividad del término.
- Precisión científica y técnica del término.

### **6.2.3 Establecimiento de facetas y subfacetas.**

El lenguaje sintético permitió que se establecieran las facetas y subfacetas generales, a partir de esta labor, fueron ubicados los términos extraídos de las fuentes técnicas pertenecientes al lenguaje natural, lo cual permitió el desarrollo de la estructura del tesauro de lo general a lo específico.

Por lo anterior, el tesoro se encuentra compuesto por las siguientes facetas:

## HIDROCARBUROS

ALMACENAMIENTO DE HIDROCARBUROS

ASOCIACIÓN EN HIDROCARBUROS [+]

CARGA DE HIDROCARBUROS [+]

COMERCIALIZACIÓN DE HIDROCARBUROS [+]

CONTAMINACIÓN POR HIDROCARBUROS

DERECHO DE LOS HIDROCARBUROS [+]

DERIVADOS DE LOS HIDROCARBUROS [+]

DERRAME DE HIDROCARBUROS

DISTRIBUCIÓN DE LOS HIDROCARBUROS [+]

EMPRESA DE HIDROCARBUROS [+]

EXPLORACIÓN DE HIDROCARBUROS [+]

EXPLOTACIÓN DE HIDROCARBUROS [+]

EXPORTACIÓN DE HIDROCARBUROS

EXPROPIACIÓN DE HIDROCARBUROS

INDUSTRIA DE LOS HIDROCARBUROS

INFORMACIÓN DE HIDROCARBUROS [+]

MAQUINARIA PARA HIDROCARBUROS [+]

MEDICIÓN DE HIDROCARBUROS

NORMAS TÉCNICAS DE LOS HIDROCARBUROS

PRECIO DE LOS HIDROCARBUROS

PROCESAMIENTO DE HIDROCARBUROS

PRODUCTOR DE HIDROCARBUROS [+]

[PROPIEDADES FÍSICAS DE LOS HIDROCARBUROS](#)

[REFINACIÓN DE HIDROCARBUROS](#) [+]

[REGALÍAS DE HIDROCARBUROS](#)

[RELACIONES DE LOS HIDROCARBUROS](#) [+]

[RESERVAS DE HIDROCARBUROS](#)

[SATURACIÓN DE HIDROCARBUROS](#)

[SERVICIO TÉCNICO EN HIDROCARBUROS](#)

[SERVIDUMBRE DE HIDROCARBUROS](#)

[TIPOS DE HIDROCARBUROS](#) [+]

[TRANSPORTE DE HIDROCARBUROS](#) [+]

[YACIMIENTO DE HIDROCARBUROS](#) [+]

#### **6.2.4 Validación de términos.**

Para facilitar el proceso de validación de los términos, fue necesario transcribir la información acumulada en los formatos de recolección de términos en soporte papel y migrarlos con sus respectivas relaciones a una hoja de excel<sup>59</sup>, esto con el propósito de agilizar el tratamiento de cada uno de los términos y proveer de un esquema taxonómico del tesoro que facilitara la visualización de los términos por parte del grupo interdisciplinar que coopero en este proyecto.

Las fuentes bibliográficas especializadas para la construcción de tesauros, recomiendan que la validación debe llevarse a cabo mediante la consulta de fuentes técnicas y con la participación de un grupo interdisciplinar que avale la estructura del tesoro.

---

<sup>59</sup> Excel es un software de Microsoft Office que le permite crear tablas, calcular y analizar datos.

Las fuentes técnicas consultadas fueron:

- Diccionarios especializados.
- Vocabulario controlado especializado.
- Sitios web especializados en el tema.
- Repertorio legislativo sobre el tema.
- Documentos técnicos emitidos por organismos en la materia.
- Libros especializados en hidrocarburos en el contexto latinoamericano.

Para acatar la metodología propuesta de validación, se conformó un grupo interdisciplinar de profesionales que cuentan con elementos teóricos y prácticos sobre el tema a desarrollar, se contó con la colaboración de ingenieros de petróleos pertenecientes a los centros de información técnica como fue el caso de Ecopetrol y el Ministerio de Minas y Energía y el especialista en lenguajes documentales y docente de la Universidad de la Salle, el profesor Edwin Alonso Sossa Cruz.

En lo que respecta a la validación de términos en forma y composición, el Director de trabajo de grado contribuyó significativamente en este proceso, debido a que cuenta con un amplio conocimiento en construcción de tesauro.

Producto de la metodología empleada para la validación, éstos fueron los resultados del tesauro en relación a la cantidad de términos y establecimiento de relaciones:

- 1.130 términos descriptores,
- 730 relaciones entre términos

- 113 términos equivalentes y
- 12 notas de alcance

Desde la fase inicial de extracción hasta la fase de validación fueron depurados 1507 términos, debido a que estas unidades lingüísticas no correspondían temática al objeto que motivo la construcción del tesoro, presentaban ambigüedades conceptuales y/o no enriquecían terminológicamente a éste.

### **6.2.5 Criterios para la construcción de tesauros.**

Las facetas y subfacetas cuentan con diferentes niveles de amplitud y especificidad, éstos son proporcionales a la riqueza terminológica del contenido registrado en los documentos técnicos generados en esta área del conocimiento.

Para nuestro caso, los criterios preestablecidos que describiremos a continuación, deben ser tenidos en cuenta para la estructuración, mantenimiento y actualización de un tesoro homogeneizado.

- **Criterio de Transmutabilidad**

Este criterio reconoce la adaptación, reinención y modificación de las ciencias y por consiguiente de su conocimiento. Estos cambios, innovaciones o ajustes del conocimiento se evidencian, reflejan, representan y nominan a partir de las palabras que emplean.

En la construcción de tesauros, es importante reconocer que los términos siendo susceptibles de adaptación, cambio y evolución, deben ser identificados y evidenciados en el mismo sistema. Ello hace necesario que las personas a cargo de la gestión de tesauros se encuentren en continua investigación con el fin de evitar la desactualización de la herramienta y la ambigüedad en el empleo de las palabras del tesoro,



ante el uso o desuso de los mismos por la ciencia o disciplina que los genera.

Para controlar el criterio anteriormente enunciado, es recomendable hacer uso de las diferentes Notas aplicables a los tesauros, con el fin de reducir la ambigüedad de los términos, se hace entonces necesario indicar las condiciones de uso del término cuando la situación a describir lo requiera, con el fin de romper cualquier tipo de ambigüedad lingüística.

Este tipo de control debe centrarse en reflejar el proceso evolutivo de las palabras y sus significados, a partir del tiempo o los contextos de aplicación.

- **Criterio Estructural.**

La arquitectura de un tesoro debe evidenciar una estructura sistemática, este quiere decir que su estructura debe ser planeada y organizada como un todo.

La estructura sistemática de un tesoro se logra por la aplicación de los 3 tipos de relaciones básicas que lo caracterizan, estos tres tipos de relaciones son:

- La relación jerárquica: Como su nombre lo indica, establece los niveles de subordinación entre los términos y nos permite identificar los aspectos concretos y específicos sobre los que se posee información relacionada con un fenómeno, sujeto o hecho jurídico específico. Este tipo de relación suministra conocimiento a los usuarios en la medida que les permite visualizar y desglosar aquellos aspectos referentes a un tema.

Este tipo de relación permite efectuar agrupaciones a partir de las propiedades o características del tema o fenómeno.

- La relación de equivalencia: La riqueza del idioma nos brinda múltiples formas de representar y nominar un tema o asunto, pero por principio normalizador y de homogenización solo una debe considerarse como forma aceptada y las otras como formas alternativas que nos facilitaran el acceso o consulta.

Este tipo de relación nos permite enriquecer el tesoro con el fin de garantizar diversidad terminológica a la hora de buscar o recuperar un fenómeno jurídico y sirve para indicarle al usuario el término correcto adoptado por normalización pero sin descartar las otras formas como mecanismo de consulta y acceso. Este tipo de relación guía al usuario en sus búsquedas a partir del tesoro.

- La relación asociativa: Esta relación es la que establece reciprocidad y conexión con otros términos con los que maneja complementariedad o similitud, sirve para hacer visible al usuario otros posibles temas de complementación sobre el tema que está buscando.

Es importante recalcar que las relaciones anteriormente mencionadas son relaciones reciprocas, es decir, que se dan en ambos sentidos y sirven para ordenar, agrupar, controlar y asociar las palabras contenidas en un tesoro, permitiendo que la estructura del tesoro sea manejable, solida, usable y entendible y que sirva de guía para la consulta o navegación por parte tanto de los abogados como de los usuarios en general.

Este Criterio de estructuración va a permitir que al hacer una búsqueda se sugiera el término empleado y normalizado de forma adecuada y se identifique sus equivalentes u otros términos con los que maneja similitud o complementariedad.

- **Criterio de Especialización del término.**

Los términos deben de tener sentido dentro del área o disciplina que los crea y los define, estableciendo relación con términos afines desde

otras disciplinas, pero respetando su forma habitual de uso desde la especialidad que los trata.

- **Criterio de profundidad del término:**

Se hace necesario que a la hora de construir los términos del tesoro estos se construyan lo más completos posibles (pero sin abusar de su extensión), con el fin de lograr mayor precisión y exactitud a la hora de nombrar los temas o fenómenos jurídicos representados en las sentencias.

Bajo este criterio, se debe tener mucho cuidado de no caer en la construcción de términos inexistentes, redundantes, inconclusos, incompletos o sin sentido, ya que desvirtúan la representación temática de los contenidos.

Resulta recomendable no excederse en la asignación o empleo de palabras para la construcción de términos compuestos, por lo que se hace necesario que la descripción y representación de los fenómenos jurídicos sea concreta.

- **Criterio semántico:**

Una correcta construcción del término exige una adecuada construcción gramatical y conexión entre las palabras, permitiéndonos identificar fenómenos jurídicos específicos y concretos sin ambigüedad, por lo tanto la construcción de los términos del tesoro requiere la definición de unas políticas semánticas y gramaticales que orienten su adecuada construcción.

Para este tipo de tesauros, es recomendable que los términos sean construidos bajo una forma directa, es decir, no es recomendable emplear procesos de construcción invertida (uso de guiones), ya que esa forma de construcción de términos era empleada anteriormente solo para la normalización en catálogos manuales o de fichas catalográficas.

La capacidad de indexación, organización y procesamiento de las tecnologías actuales, modificó sustancialmente la práctica de construcción de los tesauros y por lo tanto en la actualidad no se aplican los guiones.

El establecimiento de estas reglas semánticas, conllevan a lo siguiente:

- Adecuada estructura gramatical del término: El término debe permitir una adecuada representación del fenómeno jurídico que se quiere categorizar, lo que lleva a evaluar las distintas connotaciones que ha tenido en el tiempo. Eso garantiza calidad del proceso, porque si se selecciona mal el término viene mala la titulación e identificación de la línea.
- Precisión: El término debe representar con exactitud el fenómeno jurídico que debemos evidenciar o nominar, lo que implicará en algunas ocasiones efectuar análisis conceptuales para lograr mayor precisión en el uso y aplicación de los términos.
- Propiedad, concisión y uso: Los términos deben ser un reflejo del uso que hacen de ellos sus usuarios, por consiguiente, para que el tesoro sea efectivo es necesario conocer la terminología que se emplea y darle preponderancia a esta sobre otras formas.

La normalización no implica que se deba desconocer la estructura de los términos, antes por el contrario, debe definir y preestablecer una forma adecuada y aceptada para su uso.

### **6.2.6 Reglas semánticas del Tesoro.**

Para garantizar uniformidad en la construcción del tesoro se conformaron las siguientes reglas semánticas:

**A. Descriptor:** Constituye en el término que debe ser utilizado para la indización, que representan temáticamente con precisión el contenido de los documentos y en las solicitudes de recuperación de información.

**i. Tipos de términos descriptores:**

- **Descriptor simple:** Es aquel conformado por una sola palabra y/o representa una sola unidad de pensamiento.

Ejemplo: Sismo

- **Descriptor compuesto:** Son términos compuestos por varias palabras que se caracterizan porque indican mayor especificidad, precisión y profundidad temática, además agrupa y representa mayor información para las búsquedas.

**B. No descriptor:** Es el término que no debe usarse para la indización de documentos, debido a que constituye en términos sinónimos que no fueron validados como términos descriptores por el grupo de especialistas y las fuentes especializadas. Estos deben incluirse para abarcar las posibles entradas utilizadas por el usuario en la recuperación de la información.

En el aplicativo de TemaTres se encuentran en minúscula y precedidos en de la abreviatura UP.

**C. Forma de presentación del término descriptor:**

- **Singular:** Siempre debe procurarse el uso de términos descriptores en singular.

- **Plural:** Cuando es un concepto abstracto está representado por más de un miembro es recomendado utilizar el plural.
- **Sustantivos:** Cuando las palabras en singular y plural tengan diferente significado o cuando se quiera diferenciar una técnica de su producto.
- **Adjetivos:** Debe evitarse el uso aislados de adjetivos; en caso de requerirse por su utilización en los usuarios, es necesario crear una referencia cruzada que asocie el término en forma adjetiva a términos compuestos.
- **Adverbios:** No son admitidos, sin embargo cuando la expresión técnica del término lo requiere, se debe incorporar.
- **Verbos:** No se admiten verbos participativos o infinitivos y deben convertirse a términos sustantivados.
- **Abreviaturas, siglas y acrónimos:** Debe evitarse el uso de abreviaturas, sigla y acrónimos como términos descriptores, no obstante cuando éstos sean de mayor recordación por el usuario, es importante crear la referencia reciproca desde la forma abreviada.
- **Ortografía:** En caso de existir más de una variación ortográfica de un término, todas deben incluirse en el tesauro, debido a que son puntos de acceso que guían al usuario hacia el término descriptor.  
El término que tenga mayor uso dentro de la literatura se establecerá como término descriptor y las demás formas ortográficas del término es necesario relacionarlos mediante el uso del símbolo (UP) para que se identifiquen como términos no preferentes.

- **Extranjerismos:** Es necesario pero no obligatorio incluir los términos en otro idioma y pueden ser términos descriptores o términos equivalentes según la frecuencia del uso en el concepto según el usuario.
- **Denominaciones coloquiales y nombres científicos:** Cuando la representación de un concepto es de mayor uso dentro de la jerga popular que científico, se debe procurar su uso coloquial y es necesario establecer relación de equivalencia con el término científico.
- **Nombres de lugares:** Los nombre de los lugares varían de idioma a idioma, por lo que es necesario utilizar como término descriptor el de mayor uso entre los usuarios.
- **Género:** Los términos deben ser preferiblemente en género masculino siempre y cuando no modifique su significado.

## 6.3 AUTOMATIZACIÓN DEL TESAURO

### 6.3.1 Herramienta para la automatización del tesauro.

En la actualidad, la construcción de tesauros es apoyada con la utilización de aplicaciones y software especializado que automatiza las actividades que inicialmente se ejecutaban manualmente, en gran medida este cambio se debe a la migración de los tesauros de soportes físicos a ambientes electrónicos, lo cual ha facilitado su acceso y genera mayor interacción para el desarrollo y

administración del tesoro mediante el trabajo cooperativo apoyado en el uso de las tecnologías de comunicación como la Internet.

Teniendo en cuenta que la construcción de éste tesoro electrónico constituye en un elemento que facilita la participación, el intercambio de información, acceso y desarrollo terminológico en temas relacionados con hidrocarburos, mediante la participación de los diferentes actores como entidades públicas, privadas y la comunidad en general, es vital que su disposición y publicación se encuentre bajo la política de libre acceso.

El tesoro no cuenta con restricciones de uso y desarrollo (sin que eso implique que no exista control), por lo anterior es necesario que todos los elementos que integran el software que automatiza el tesoro, deban estar ligados a esta corriente de libre acceso.

La licencia GPL (Licencia Pública General)<sup>60</sup>, es un elemento que brinda la posibilidad de hacer uso del software para la construcción del tesoro sin incurrir en violaciones de derechos de autor.

Para éste proyecto es necesario contar con un software de administración de bases de datos que es el encargado de almacenar los datos bajo un modelo Entidad – Relación y un administrador de tesoros, el cual tiene por función de efectuar procesos de ingreso, actualización, consulta y eliminación de datos (términos), siendo éste el que presenta la interfaz y el contacto directo con el administrador y los usuarios.

---

<sup>60</sup> GENERAL PUBLIC LICENSE: Licencia de regulación de los derechos de autor de los programas de software libre (free software) la cual es promovida por la Free Software Foundation (FSF) en el marco de la iniciativa GNU.



Los criterios de selección para el software libre de gestión de tesauros son:

- Permitir la aplicación de las relaciones semánticas existentes entre términos.
- Permitir la navegación asociativa y jerárquica para facilitar la ubicación de un término en específico.
- Contar con protocolos de intercambio de información como XML (eXtensible Markup Language)<sup>61</sup> y en RDF (Resource Description Framework)<sup>62</sup>.
- Disponer de un área de búsquedas, donde se ingrese el término a localizar y presentarlo con las relaciones semánticas existentes.
- Disponer de administración de perfiles para establecer control en el uso del tesoro.
- Desarrollo en un ambiente PHP<sup>63</sup> que es un lenguaje GNU<sup>64</sup> lo que permite que opere eficientemente sobre bases de datos relacionales como MySql, Appserver ya sea en sistemas operativos Windows ó Linux.

---

<sup>61</sup> XML: Un lenguaje de marcado extensible que puede usarse para almacenar datos en un formato estructurado, basado en texto y definido por el usuario.

<sup>62</sup> RDF: Es una infraestructura general para describir los metadatos de un sitio Web, ó la información acerca de la información en dicho sitio. Provee interoperabilidad entre las aplicaciones que intercambian información en lenguaje de máquina a través de la Web. El RDF detalla información tal como el mapa de un sitio, las fechas de actualizaciones, palabras clave para los buscadores y derechos de propiedad intelectual de la página.

<sup>63</sup> PHP: Es un lenguaje de programación usado generalmente para la creación de contenido para sitios web. El nombre es el acrónimo recursivo de "PHP: Hypertext Preprocessor" (inicialmente PHP Tools, o, Personal Home Page Tools), y se trata de un lenguaje interpretado usado para la creación de aplicaciones para servidores, ó creación de contenido dinámico para sitios web.

<sup>64</sup> GNU: Proyecto creado en 1984 con el fin de desarrollar un sistema operativo tipo UNIX según la filosofía del "software libre".

Para la automatización del tesoro se evaluaron los siguientes softwares:

- TemaTres
- TheSage´s English Dictionary and Thesaurus
- Term Tree

Para la evolución se elaboró y utilizó el “Formato para la evaluación de software para la administración de tesauros”, el cual puede ser consultado en el Anexo B, que cuentan con elementos que permiten identificar el software que más se ajusta a las necesidades del proyecto.

Producto de la evaluación del aplicativo, fue seleccionado *TemaTres*, debido a que presenta las siguientes cualidades diferenciadoras, que lo aventajaron sobre los demás software.

- Es libre y su código fuente es abierto.
- Utiliza como base de datos relacional MySQL, aplicativo GPL (software libre).
- Gran capacidad de almacenamiento de términos en la base de datos (MySQL), y se encuentra limitada no por software sino por la capacidad del hardware.
- Se efectúan actualizaciones del software desde el creador y su migración es automática.
- Establece las relaciones jerárquicas, de equivalencia y alfabéticas (asociativas).
- Controla la sinonimia y establece relaciones entre términos descriptores y no descriptores.
- Permite descargar los términos en múltiples formatos para hacer uso de otras herramientas.

- La velocidad para el ingreso, búsqueda, establecer relaciones, modificación y eliminación de términos depende del ancho de banda del computador y no del software.
- Permite el control de metadatos utilizando MADS<sup>65</sup>, XML, SKOS-Core<sup>66</sup>, XTM<sup>67</sup>, DC<sup>68</sup>.
- Los términos cuentan con hipervínculos que permiten navegar entre el tesoro.
- Cuenta con links sobre los términos que remiten a páginas Web con el fin de ampliar información sobre el término seleccionado, las páginas referenciadas son Voced database, google book, google images, google scholar, google y wikipedia. Si se desea, permite modificar éstos vínculos a otras páginas externas del tesoro, de acuerdo a la necesidad del administrador del tesoro.
- Permite la publicación en Internet e Intranet.
- Permite la creación de perfiles para búsqueda, establecer relaciones, modificación y eliminación según el caso.
- Interfaz sencilla e intuitiva para usuarios y administradores.

El tesoro desarrollado puede ser consultado en la siguiente dirección:

<http://www.knoc.com.ar/hidrocarburos/index.php>

---

<sup>65</sup> MADS: Esquema de metadatos de descripción y autoridad basado en XML.

<sup>66</sup> SKOS-Core: Web semántica avanzada.

<sup>67</sup> XTM: Es la fusión entre XML y topic maps, como medio de intercambio de metadatos.

<sup>68</sup> DC: Dublin Core es una organización abierta y libre para el desarrollo de estándares de metadatos interoperables en línea.

### 6.3.2 Presentación del tesoro.

Teniendo en cuenta que el tesoro está dirigido a un lugar geográfico particular dentro del territorio Colombiano, la utilización símbolos fueron extraídos de la Norma Técnica Colombiana NTC 1476: 1998, debido a su compatibilidad con la norma española UNE 50-106-90.

En la siguiente tabla se presentan los símbolos, descripción y definición tomada de la NTC 1476:1998.

**Tabla 20. Símbolo, descripción y definición de la NTC 1476:1998.**

ABREVIATURA	DESCRIPCIÓN	DEFINICIÓN NTC 1476:1998
TG	Término Genérico	El término que sigue representa un concepto con la connotación más amplia.
TE	Término Especifico	El término que sigue se refiere a un concepto con una connotación más precisa.
TR*	Término Relacionado	El término que sigue está asociado pero no es un sinónimo, cuasi-sinónimo, término genérico o término específico.
NA	Nota de Alcance	Una nota que se adjunta a un término para indicar su significado restringido dentro de un lenguaje de indización.
USE	Término preferido	El término que sigue es el término preferido cuando se debe elegir entre sinónimos o cuasi-sinónimos.

Fuente: El autor.

TemaTres es compatible con el uso de los símbolos de la Norma Técnica Colombiana, lo que establece la correcta presentación de las relaciones entre términos del tesoro.

Igualmente TemaTres indica la cantidad de subordinaciones de un término descriptor, conforme se introduzca en la estructura del tesoro, por lo que una medida válida es la asignación de un número posterior al símbolo (TE), porque representa el nivel de amplitud de la faceta-término que se está consultando en el tesoro.

A medida que se profundice temáticamente en la faceta, el número debe aumentar y en caso contrario disminuir.

### **6.3.3 Presentación jerárquica.**

Una de las particularidades del software TemaTres, es que nombra la *presentación jerárquica* como se identifica en las fuentes bibliográficas consultadas para la construcción de tesoro, bajo el nombre de *presentación sistémica*, sin embargo acuña perfectamente función.

Ésta consiste en la presentación de los términos de acuerdo a la jerarquía establecida en la superordenación a subordinación al interior de las facetas. Cada término descriptor presenta sus relaciones con términos específicos (TE), términos relacionados (TR) y los términos que no deben ser usados para la indización de documentos representados por los símbolos (USE/UP).

En la interfaz de inicio aparecen las tres principales facetas que componen el tesoro.

#### **6.3.4 Presentación alfabética.**

Es la presentación del universo de términos expresados de manera alfabética, en el cual se crean puntos de acceso sobre los términos descriptores para el indizador o el usuario.

#### **6.3.5 Actualización y modificaciones.**

Todo tesoro es susceptible de actualización y modificación, por lo que se ha determinado que el tesoro, se actualizará y modificará una vez se registre cambios terminológicos que se presentan dentro del dominio, para lo cual es necesario el uso del “Formato para extracción y actualización del tesoro en hidrocarburos” (Véase Anexo A) para generar registros que permitan la trazabilidad de la actividad.

Por tratarse de un tesoro perteneciente a un dominio especializado la necesidad de revisión se acentúa, debido a que cuanto más especializada es la terminología, existe mayor probabilidad que los términos preferentes de indización se vuelvan obsoletos rápidamente.

Como Lancaster indica “un vocabulario controlado no puede ser estático; tiene que crecer. Un vocabulario elaborado a partir de términos extraídos de la literatura crecerá de forma muy rápida al principio, para ir decreciendo progresivamente. Su amplitud dependerá no solo del campo temático sino también de la especificidad y el tipo de términos”<sup>69</sup> y complementa “el crecimiento del vocabulario debe ser

---

<sup>69</sup> LANCASTER. P. 125.

continuo y basarse en la autoridad literaria y en la del usuario: los nuevos términos se obtendrán a partir de operaciones de indización y recuperación. Serán términos necesarios para expresar aspectos temáticos de los documentos y las búsquedas y que no puedan ser representados por los términos existentes en el vocabulario. Este es el criterio fundamental que debe seguirse para incorporar nuevos términos al vocabulario”<sup>70</sup>.

Se tendrán en cuenta los siguientes argumentos para realizar la actualización:

- Incorporación de la terminología derivada del grado de evolución del lenguaje especializado.
- La opinión de los especialistas en la materia.
- Las consultas de los usuarios.
- La utilización de los descriptores en el proceso de indización.

### **Características de los posibles cambios.**

- a. En la actividad de indización aparecen nuevos términos y estos se listarán como posibles términos que harán parte del tesoro, su inclusión depende del grado de aparición y/o frecuencia de uso. Esta actividad se hará con el apoyo de un especialista para que asigne la ubicación del término en el tesoro junto con sus relaciones.
- b. Cambio de términos descriptores a término no descriptor por tratarse de términos que dentro de las relaciones jerárquicas no corresponden a un término preferente o que su frecuencia de uso es mínima, sin embargo no

---

<sup>70</sup> *Ibíd.*, P.127.

se eliminan, sino que se mantendrán debido a que constituyen en punto de acceso al tesoro para los usuarios. Recuerde añadir la sigla USE.

- c. Traslado de un término descriptor a otro grupo jerárquico, debido posibles fallas de ubicación en la construcción del tesoro.



## 7 CONCLUSIONES

Este trabajo de investigación me aporó significativamente ya que me permite percibir que el rol del profesional de la información para el nuevo siglo se percibe como un arquitecto de la misma, ya que en ocasiones las técnicas de procesamiento y tratamiento de la información no pueden ser tomadas tan ceñidas a la normas, los nuevos escenarios han generado especialización en nuestra profesión, exigiéndonos mayor investigación, análisis, interacción con los usuarios, dominio de tecnologías y conocimiento de tendencias en materia de gestión de información. Solo analizando nuestra profesión desde este panorama, podremos responder a un escenario diverso y cambiante, ajustándonos y ajustando nuestra disciplina a una demanda real sobre lo que nuestras comunidades o entornos demandan.

Evidencio desde este proyecto que la arquitectura de información para que sea efectiva debe ser diseñada bajo un conocimiento, investigación y demarcación de un escenario concreto y real, los escenarios diversos y amplios no logran para el momento actual responder de forma óptima a las necesidades de información de los usuarios. Por consiguiente logrando entender y aplicar estas premisas, impulsaremos el reconocimiento y posicionamiento social de nuestra profesión como elemento motor de cambio social, bajo premisas de equidad e igual de acceso a la información.

Los tesauros son y han sido objeto de estudio de la Bibliotecología, al menos, por dos razones: porque se utilizan como instrumento para la organización del conocimiento y porque se utilizan como herramienta para la indización y la recuperación de la información.

El tesauro guarda una relación estrecha con la Terminología en la medida que es una estructura de descriptores basada en relaciones que presumiblemente se establece entre los conceptos de un documento, conceptos expresados básicamente mediante unidades terminológicas.

Para que se lleven a cabo estos proyectos de manera exitosos, el profesional de Ciencias de la información debe ser el puente que enlaza las necesidades de los usuarios con los recursos de información, mediante la utilización de elementos que brinden soluciones innovadoras.

El diseño de un tesauro de información de actualidad, a partir de los términos de los documentos y de las peticiones de los usuarios, se caracteriza, además de por la riqueza de sinónimos, por un elevado nivel de precoordinación.

La relación que existe entre la Informática y los tesauros, puede evidenciarse, primero a través de las aportaciones teóricas y metodológicas con las que la Informática contribuye al proceso de indización y la recuperación de la información en la que interviene el tesauro. Segundo, mediante la descripción, de manera general, del procedimiento básico para la elaboración de un tesauro a partir de la informática.

En fuentes bibliográficas especializadas en construcción de tesauros, se evidencia que existen divergencias conceptuales entre autores, situación que causa efectos negativos a personas no expertas, por lo que estas deben homologar conceptos necesarios para construir tesauros. Aún más parece contradictorio cuando los

tesauros suponen que su construcción es con propósitos de normalización terminológica, y se presenten incompatibilidades conceptuales.

La necesidad de dotar a los usuarios con estructuras de conocimiento que les permitan acceder de manera rápida y amigable a la información, se consigue en los tesauros debido a que conjugan a través del control terminológico la precoordinación y postcoordinación una posibilidad real a la práctica de descripción de información especializada.

La información crece exponencialmente y esto se debe en gran medida a la conectividad y los medios tecnológicos, creando de esta manera que las comunidades expertas produzcan información especializada, lo cual requiere herramientas de descripción que extraiga el contenido de sus producciones científicas mediante el uso de vocabularios controlados que facilitan la recuperación de información.

Mediante los tesauros digitales de libre acceso se fomenta la cooperación nacional e internacional para el desarrollo del recurso y a su vez puede ser fuente para la construcción de nuevos tesauros, optimización de los existentes e integración con otros desde una óptica interdisciplinar.

## 8 RECOMENDACIONES

Hasta el momento el bibliotecólogo se ha preocupado más por la revisión de las normas y no por el enriquecimiento de la metodología para la creación de tesauros. A partir del rescate de las propuestas hechas desde otras disciplinas, es posible enriquecer y replantear la forma en que se diseñan los tesauros, ya que los mecanismos, herramientas y técnicas para la representación y organización del conocimiento vienen experimentando cambios sustanciales y estructurales a partir de los avances tecnológicos y su papel central como herramientas mediadoras para la gestión de información, generando escenarios más dinámicos, enriquecedores, flexibles y ajustados a los usuarios, a partir de procesos de interacción mucho más amigables.

Las Universidades deberían de prestarle mayor atención a esta tendencia ya que los procesos de gestión y arquitectura de información se percibe como una competencia propia e intrínseca del profesional de la información para el nuevo siglo, por consiguiente es recomendable revisar y reforzar estas asignaturas al interior de nuestra formación, con el fin de ampliar los escenarios y perspectivas de nuestro rol profesional, permitiéndonos incursionar en estos nuevos escenarios de la gestión de la información y el conocimiento.

## BIBLIOGRAFÍA

### Fuentes impresas

AMERICAN NATIONAL STANDARD INSTITUTE, NATIONAL INFORMATION STANDARD ORGANIZATION. NORMA: ANSI/NISO Z39.19 - 2003. Guidelines for the construction, format and Management of monolingual thesauri. P. 56.

ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. UNE Directrices para el establecimiento y desarrollo de tesauros monolingües. Madrid: AENOR, 1990, P.48. (50-106)

ENCABEZAMIENTOS DE MATERIA: normativa para su redacción. Madrid: Biblioteca Nacional, Ministerio de Cultura, 1991. P. 25.

GIL URDICIAIN, Blanca. Manual de lenguajes documentales. Madrid : Ediciones Trea, 2004. P.280.

GIL URDICIAIN. Blanca. Orígenes y evolución de los tesauros en España. p. 63 – 110. EN : Revista General de Información y Documentación. Vol. 8, No. 1 (1998). P.342.

INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS Y CERTIFICACIÓN. Guía para el establecimiento y desarrollo de tesauros monolingües. Bogotá : ICONTEC, 1998. P.66 (NTC 1476)

INSTITUTO COLOMBIANO PARA EL FOMENTO DE LA EDUCACIÓN SUPERIOR. Aprender a investigar. Bogotá. 1999. P.130.

LANCASTER, Frederick Wilfrid. Control del vocabulario en la recuperación de información. 2 ed. Madrid: Universidad de Illinois, 1995. P. 283.

LÓPEZ, Miguel Ángel. Evolución histórica y tendencias observables en los tesauros. Departamento de Biblioteconomía y Documentación. Madrid. Universidad Carlos III de Madrid. 2006. P 370.

MARTÍNEZ BENCARDINO, Ciro. Estadística y muestreo. 11 ed. Bogotá : Ecoe Ediciones, 2003. P..879

NAUMIS PEÑA, Catalina. Los tesauros documentales y su aplicación en la información impresa, digital y multimedia. México. Universidad Autónoma de México. 2007. P.288

ORGANIZACIÓN INTERNACIONAL DE NORMALIZACIÓN. Documentation - Guidelines for the establishment and development of monolingual thesauri: ISO, 1986, P. 32. (ISO 2788)

RODRÍGUEZ DELGADO. Rafael. Teoría de sistemas y tesauros facetados. En: Revista Española de Documentación Científica, vol. 1, No. 3. (Dic., 2005) p. 213-224.

SLYPE, Georges Van. Los lenguajes de indización: concepción, construcción y utilización de los sistemas documentales. Madrid : Fundación Germán Sánchez Ruipérez, 1991. P. 198.

## Fuentes electrónicas

ARANO, Silvia. Los tesauros y las ontologías en la Biblioteconomía y la Documentación. [Termino de búsqueda: tesoro y taxonomías]. [En línea-html]. [consulta: 12 de nov., de 2014]. Disponible en: <<http://www.hipertext.net/web/pag260.htm>>

CRUZ RUBIO. María. Análisis documental: indización y resúmenes en bases de datos especializadas. [En línea]. Madrid : CINDOC, 2000. P. 15. [consulta: 14 de ago., de 2014]. Disponible en: <<http://www.iberius.org/es/AisManager?Action=ViewDoc&Location=getdocs:///DocMapCSDOCS.dPortal/2519>>

ESCUELA UNIVERSITARIA DE BIBLIOTECOLOGÍA Y CIENCIAS AFINES. Diccionario de organización y representación de conocimiento, clasificación, indización, terminología. [Termino de búsqueda: leguajes documentales]. [En línea-html]. Uruguay.[consulta: 25 ago., 2014]. Disponible en: <[http://www.eubca.edu.uy /diccionario/letra\\_l.htm](http://www.eubca.edu.uy /diccionario/letra_l.htm)>

LAMARCA LAPUENTE, María Jesús. Hipertexto, el nuevo concepto de documento en la cultura de la imagen. [En línea]. Madrid : Universidad Complutense de Madrid. [consulta: 10 de ago., de 2014]. Disponible en: <[www.hipertexto.info/documentos/xfml.htm](http://www.hipertexto.info/documentos/xfml.htm)>

LEMB. Lista de Encabezamiento de Materias para Bibliotecas. [Término de búsqueda: LEMB], [en línea - html]. [consulta: 25 de ago., de 2014]. Disponible en: <[www.accessmylibrary.com/coms2/summary\\_0286-32034427\\_ITM](http://www.accessmylibrary.com/coms2/summary_0286-32034427_ITM)>

LÓPEZ-HUERTAS PÉREZ, María José. EN : Boletín de ANABAD, Tomo 41, No. 1, 1991. [en línea -PDF]. Madrid, 1991. P 61 – 79. [consultado: 10 de ago., de 2014]. Disponible en: <<http://dialnet.unirioja.es/servlet/extaut?codigo=267457>>  
[ISSN 0210-4164](http://dialnet.unirioja.es/servlet/extaut?codigo=267457)>

MONTGOMERY COUNTRY MEMORIAL LIBRARY SYSTEM. Sistema de Clasificación Decimal Dewey. [termino de búsqueda: SCDD], [en línea - html]. [consulta: 25 de ago., de 2014]. Disponible en :<[www.countylibrary.org/span3.htm](http://www.countylibrary.org/span3.htm)>

R020 BIBLIOTECOLOGÍA Y CIENCIAS DE LA INFORMACIÓN - TEMATRES. Software libre para la gestión de tesauros, 2008. [termino de búsqueda: terminología en desastres], [en línea – html] Buenos Aires. [consulta: 30 de nov., de 2014]. Disponible en: <<http://www.r020.com.ar/tematres/>>

ROGET, Peter Mark. Thesaurus of English words and Phrases Classified and Arranged so as to Facilitate the Expression of Ideas and Assist in Literary Composition. [termino de búsqueda: Roget]. [en línea-html]. [consulta: 12 de nov., de 2014]. Disponible en: <<http://www.bartleby.com/110/>>

THE LIBRARY OF CONGRESS. Library of Congress Classification. [termino de búsqueda: LCC], [en línea - html]. [consulta: 25 de ago., de 2014]. Disponible en: <<http://www.loc.gov/catdir/cpsol/lcc.html>>

VOCABULARY SERVER BLOG. [en línea]. [termino de búsqueda: tematres], [en línea - html] Buenos Aires. [consulta: 30 de nov., de 2014]. Disponible en: <<http://www.vocabularyserver.com/>>



## ANEXO A

### FORMATO DE EXTRACCIÓN Y ACTUALIZACIÓN DE TÉRMINOS PARA EL TESAURO EN HIDROCARBUROS.

FORMATO PARA EXTRACCIÓN Y ACTUALIZACIÓN DE TÉRMINOS			
ANALISTA	<input style="width: 95%;" type="text"/>		
FECHA	<input style="width: 90%;" type="text"/>	<input style="width: 90%;" type="text"/>	<input style="width: 90%;" type="text"/>
	DD	MM	AAAA
	ACTUALIZACIÓN <input style="width: 40%;" type="text"/>		ADICIÓN <input style="width: 40%;" type="text"/>
			SUPRESIÓN <input style="width: 40%;" type="text"/>
			CORRECCIÓN <input style="width: 40%;" type="text"/>
TÉRMINO	<input style="width: 95%;" type="text"/>	FACETA	<input style="width: 95%;" type="text"/>
TÉRMINO GENERAL (TG)	<input style="width: 95%;" type="text"/>		
TÉRMINO ESPECÍFICO (TE)	<input style="width: 95%;" type="text"/>		
TÉRMINO RELACIONADO (TR)	<input style="width: 95%;" type="text"/>		
USADO POR (UP)	<input style="width: 95%;" type="text"/>		
USE (USE)	<input style="width: 95%;" type="text"/>		
NOTA DE ALCANCE (NA)	<input style="width: 95%;" type="text"/>		
Marque con una "X" el documento técnico del cual se extrajo el TÉRMINO	Diagnóstico Técnico	<input style="width: 40%;" type="text"/>	Nº <input style="width: 40%;" type="text"/>
	Concepto Técnico	<input style="width: 40%;" type="text"/>	Nº <input style="width: 40%;" type="text"/>
	Estudio Técnico	<input style="width: 40%;" type="text"/>	Nº <input style="width: 40%;" type="text"/>
	Otro documento	<input style="width: 100%;" type="text"/>	
FUENTES DE VALIDACIÓN:	<input style="width: 95%;" type="text"/>		
OBSERVACIONES	<input style="width: 95%;" type="text"/>		
DEFINICIÓN	<input style="width: 95%;" type="text"/>		
FECHA DE APROBACIÓN	<input style="width: 90%;" type="text"/>		FECHA DE INCORPORACIÓN
	DD/MM/AAAA		DD/MM/AAAA
RESPONSABLE	<input style="width: 95%;" type="text"/>		

## ANEXO B

### FORMATO PARA LA EVALUACIÓN DE SOFTWARE PARA LA AUTOMATIZACIÓN DE TESAUROS

SOFTWARE PARA LA AUTOMATIZACIÓN DEL TESAURO

MODULO	Ítem	ESPECIFICACIÓN TÉCNICA	TEMATRES					THE SAGE'S ENGLISH DICTIONARY AND TESAURUS					TERM TREE				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Construcción y mantenimiento	1	Es un software libre?					X					X					X
	2	Utiliza la administrador de Base de datos de software libre?					X	X					X				
	3	Permite efectuar migraciones de datos en múltiples formatos?					X			X					X		
	4	Permite la inclusión de términos ilimitadamente?					X		X					X			
	5	El ingreso de nuevos términos dependen de la capacidad de almacenamiento de la base de datos?					X					X					X
	6	Permite publicar el directorio en ambiente Web o intranet?					X	X					X				
	7	El software se encuentra en una versión actualizada?					X				X		X				
	8	Permite actualizaciones desde el desarrollador?					X			X					X		
Especificaciones técnicas	9	Establece relaciones Jerárquicas de equivalencia y asociativas entre los Términos?					X					X					X
	10	Controla la Sinonimia entre los términos y establece relaciones cruzadas entre los Descriptores y no Descriptores					X					X					X
	11	Genera Índices Sistemáticos y Alfabéticos					X					X					X
	12	Registra cambios de términos?					X					X					X
	13	Permite generar notas de alcance?					X					X					X
	14	Identifica y reporta la duplicidad de términos?					X					X					X
	15	Cuenta con búsquedas simples y avanzadas?			X							X			X		
	16	Utiliza los símbolos recomendados por las normas técnicas de construcción de tesaurus monolingües en idioma español?					X	X					X				
	17	La velocidad en el procesamiento de las búsquedas es?				X				X							
Usabilidad	18	Permite establecer hipervínculo a los documentos afines al término?					X			X			X				
	19	Posee una presentación clara y comprensible para el usuario?				X				X						X	
	20	Presenta consistencia entre los datos ingresados y su estructura?					X					X			X		
Seguridad	21	Cuenta con perfil administrador del sistema y usuario?					X					X		X			
<b>OBSERVACIONES:</b> Debido a las múltiples ventajas ofrecidas por TemaTres y producto de las ponderaciones, se decidió considerar la incorporación de la estructura del directorio en esta aplicación, porque cumple con las necesidades del Centro de Documentación e Información de esta entidad pública Distrital.			0	0	3	8	90	3	2	15	1	55	4	6	12	12	35
			96%					72%					66%				