

Universidad de La Salle

Ciencia Unisalle

Sistemas de Información, Bibliotecología y
Archivística

Departamento de Estudios de Información

10-5-2015

Construcción de un tesoro especializado en el área comercial

Henry Castro Roa

Universidad de La Salle, Bogotá

Follow this and additional works at: https://ciencia.lasalle.edu.co/sistemas_informacion_documentacion



Part of the [Library and Information Science Commons](#)

Citación recomendada

Castro Roa, H. (2015). Construcción de un tesoro especializado en el área comercial. Retrieved from https://ciencia.lasalle.edu.co/sistemas_informacion_documentacion/14

This Trabajo de grado - Pregrado is brought to you for free and open access by the Departamento de Estudios de Información at Ciencia Unisalle. It has been accepted for inclusion in Sistemas de Información, Bibliotecología y Archivística by an authorized administrator of Ciencia Unisalle. For more information, please contact ciencia@lasalle.edu.co.

**CONSTRUCCIÓN DE UN TESAURO ESPECIALIZADO EN EL ÁREA
COMERCIAL**

HENRY CASTRO ROA

**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 EL ÁREA
COMERCIAL**

HENRY CASTRO ROA

**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., octubre 5 de 2015.

AGRADECIMIENTOS

“Mi gratitud se la debo a DIOS por haberme acompañado en todos los momentos para no desfallecer y culminar una meta más en mi vida”.

“Mil gracias profesor Edwin Sossa y cuerpo de docentes de la Universidad de La Salle, por orientarme y haberme compartido sus conocimientos”.

CONTENIDO

	pág.
INTRODUCCIÓN.....	9
1 DEFINICIÓN DEL PROBLEMA	13
1.1 ANTECEDENTES DEL PROBLEMA.....	13
2. OBJETIVOS	15
2.1 OBJETIVO GENERAL.....	15
2.2 OBJETIVOS ESPECÍFICOS	15
3. JUSTIFICACIÓN	16
4. DISEÑO METODOLÓGICO	19
4.1 TIPO DE INVESTIGACIÓN	19
4.2 DELIMITACIÓN Y ALCANCE.....	32
4.3 POBLACIÓN OBJETO	33
5. MARCO TEÓRICO	34
5.1 LENGUAJES DOCUMENTALES	34
5.1.1 Definición de lenguajes documentales.....	34
5.1.3 Evolución de los lenguajes documentales.....	36
5.1.4 Lenguaje sintético y lenguaje natural controlado.	43
5.1.5 Ventajas y desventajas del lenguaje natural controlado y lenguaje sintético.....	45
5.1.6 Sistemas híbridos.....	47
5.1.7 Terminología y término.....	48
5.1.8 Términos de indización.....	49
5.1.9 Términos simples y compuestos.	49
5.1.10 Términos en singular y plural.	50
5.1.11 Accidentes gramaticales: sinónimos y polisémicos.....	52
5.1.12 Forma de presentación de los términos:	53
5.1.13 Selección de términos.....	58
5.1.14 Definición de tesoro.	64
5.1.15 Historia del tesoro.	66
5.1.16 Funciones y uso del tesoro.	70
5.1.17 Construcción de tesoros.	72

5.1.18	Definición de término descriptor y no descriptor	73
5.1.19	Características de los tesauros.....	76
5.1.20	Relaciones del tesauo.....	78
5.1.21	Relaciones y aplicación de los símbolos del tesauo.	79
6.	CONSTRUCCIÓN DEL TESAURO ESPECIALIZADO EN EL ÁREA COMERCIAL	84
6.1	METODOLOGÍA PARA LA CONSTRUCCIÓN DE UN TESAURO ESPECIALIZADO EN EL ÁREA COMERCIAL	84
6.1.1	Profundidad y extensión del Tesauo especializado en el área comercial.	85
6.1.2	Fuentes terminológicas.	85
6.2	ANÁLISIS E INTERPRETACIÓN DE LA INFORMACIÓN.....	86
6.2.1	Instrumentos de recolección de información.	87
6.2.2	Selección y recolección de términos – corpus.	87
6.2.3	Establecimiento de facetas y subfacetas.	88
6.2.4	Validación de términos.	89
6.2.5	Criterios para la construcción de tesauros.	91
6.2.6	Reglas semánticas del Tesauo en el área comercial.	95
6.3	AUTOMATIZACIÓN DEL TESAURO ESPECIALIZADO EN EL AREA COMERCIAL	98
6.3.1	Herramienta para la automatización del Tesauo.....	98
7	CONCLUSIONES	109
8	RECOMENDACIONES.....	111
	BIBLIOGRAFÍA.....	111

LISTA DE TABLAS

	pág.
Tabla 1. Definiciones de lenguajes documentales	35
Tabla 2. Ventajas y desventajas de los lenguajes documentales	46
Tabla 3. Términos de indización.	50
Tabla 4. Uso del singular y plural.	53
Tabla 5. Definiciones de sinónimos y polisemias.	55
Tabla 6. Forma de presentación de los términos.	56
Tabla 7. Selección de términos.	61
Tabla 8. Definición del término tesoro	69
Tabla 9. Características del Tesoro	81
Tabla 10. Tesoros según su idioma.	83
Tabla 11. Relaciones jerárquicas .	83
Tabla 12. Nota de alcance.	86
Tabla 13. Relaciones asociativas.	87
Tabla 14. Relaciones de equivalencia.	88
Tabla 15. Relaciones y símbolos del tesoro.	89

LISTA DE ANEXOS

	pág.
ANEXO A. FORMATO DE EXTRACCIÓN Y ACTUALIZACIÓN DE TÉRMINOS PARA LA CONSTRUCCIÓN DE UN TESAURO ESPECIALIZADO EN EL ÁREA COMERCIAL	120
.	
ANEXO B. FORMATO PARA LA EVALUACIÓN DE SOFTWARE PARA LA ADMINISTRACIÓN DEL TESAURO	121

INTRODUCCIÓN

Una de las herramientas más empleadas por las unidades de información son los Tesauros, ya que este tipo de herramientas permite mejorar y optimizar los procesos de búsqueda y recuperación de información de cara al usuario, situación que solo se logra en la medida que se obtienen puntos de acceso normalizados y acordes al ámbito de aplicación del mismo.

Cada vez es más latente y alta la demanda por parte de los usuarios, de contar con instrumentos y herramientas que les permitan guiar sus búsquedas y consultas de información en un universo tan amplio y ambiguo como el que actualmente encontramos como consecuencia de la explosión de información, situación que siembra terreno para la potencialización de los Lenguajes Documentales como instrumentos de intercambio de conocimiento.

Desde esta perspectiva, el Tesauro ya no debe ser percibido únicamente como un instrumento interno de las Unidades de Información en la realización y aplicación de actividades técnicas como es el caso del análisis de información, sino que se convierte en una herramienta de comunicación y conocimiento en el proceso de interacción entre los usuarios y la unidad de información, con miras al aprovechamiento de los recursos de información allí existentes.

Por lo tanto, este trabajo que tiene como objeto la construcción de un tesauro sobre las actividades comerciales empleadas en Colombia, alcanza gran importancia en la medida que le permite a las Unidades de Información (que albergan información sobre la materia), establecer procesos de comunicación e interacción más eficientes con sus usuarios en función de la satisfacción de sus necesidades.

En este sentido, la construcción de un tesauro especializado en el área comercial, surge de la necesidad de controlar, normalizar y homogenizar la terminología de

esta temática, permitiéndonos como profesionales de la información una descripción y representación mucho más pertinente y acertada de cara al usuario. Dicha herramienta, contribuirá a la indización de los documentos producidos en este sector o sobre esta materia, agilizará la búsqueda y recuperación de documentos y facilitará el intercambio de información entre las empresas del sector o las Unidades de Información afines a la materia.

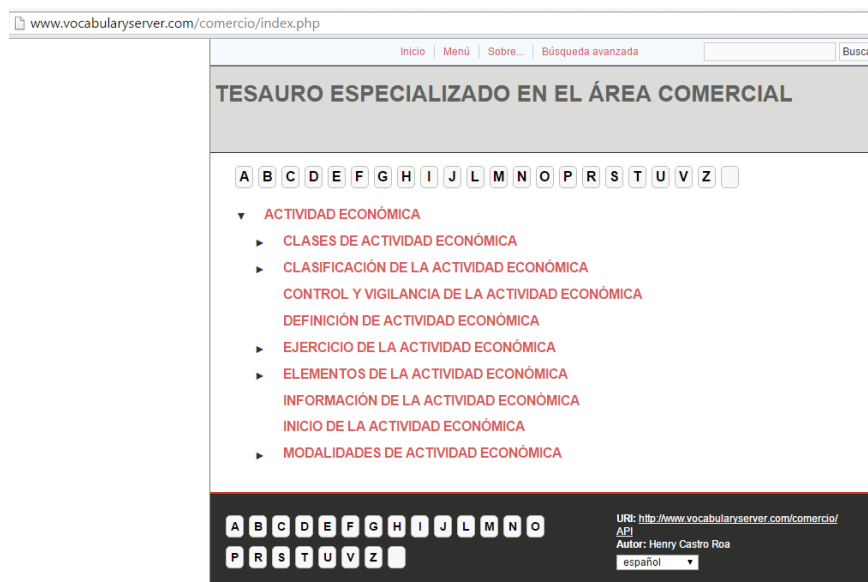
Para la elaboración del Tesauro especializado en el área comercial, se emplearon como directrices las siguientes normas:

1. La Norma ISO 25964 - el estándar internacional para tesauros y la interoperabilidad con otros vocabularios.
2. La Norma Técnica Colombiana NTC 1476:1998 – Guía para el establecimiento y desarrollo de tesauros monolingües de ICONTEC.
3. La norma Española UNE 50 – 106 – 90 Directrices para el establecimiento y desarrollo de tesauros monolingües.
4. La Norma ISO 704:2009: Terminology work -- Principles and methods
5. La Norma ISO / TC 37 - Terminología y otros recursos lingüísticos y de contenido.
6. La Norma ISO 860: Trabajo de Terminología - Armonización de conceptos y términos
7. La Norma ISO 29383: Políticas Terminología - Desarrollo e implementación.

Este Tesauro es la primera iniciativa para normalizar la terminología empleada en función de las actividades económicas discriminadas el contexto colombiano, está desarrollado en idioma español, bajo una estructura taxonómica facetada, publicado en el aplicativo TemaTres¹, el cual es un Gestor de Tesauros, se

¹ 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/>>

encuentra publicado en un servidor de acceso libre, disponible para su consulta en la siguiente dirección: <http://www.vocabularyserver.com/comercio/> , cuenta con una estructura facetada compuesta 780 términos descriptores, 730 relaciones entre términos, 123 términos equivalentes y 12 notas de alcance.



Tomada de: <http://www.vocabularyserver.com/comercio/index.php>. Consultada en octubre 05 de 2015.

Este trabajo está estructurado bajo 6 ejes fundamentales, estructurados en 8 capítulos para un mayor entendimiento del mismo: la parte 1 describe el planteamiento que orienta el trabajo de investigación, la parte 2 desarrolla un marco teórico que inicia con una conceptualización general sobre los lenguajes documentales, desarrollando aspectos inherentes con la terminología y su correcta construcción, para luego pasar a concretar el concepto de tesoro, su funcionalidad, características, creación y uso, en la parte 3 se desarrolla todo el aspecto metodológico empleado en la construcción del tesoro y en la parte 4 se procede a desarrollar el tema de la sistematización de tesauros, efectuando un análisis sobre algunas herramientas disponibles, en la parte 5 y 6 finalizamos con

las conclusiones y recomendaciones que estamos seguros guiarán a futuros investigadores que se inclinen por la construcción de este tipo de instrumentos.

Esta investigación está adscrita a la línea de investigación: información, conocimiento y sociedad; la sublínea Gestión de la Información y el Conocimiento, ya que aporta elementos para el estudio de la información y el conocimiento, considerándoles como un recurso estratégico para el avance social de los habitantes de una comunidad. A la par que desarrolla fundamentos teóricos para el análisis, la interpretación y la clasificación de la información y el conocimiento registrado en los documentos.

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 información relevante, factor que aumenta la probabilidad de no disponer de manera inmediata de información relevante, que sirva de fuente para la construcción de nuevos documentos, para la toma de decisiones y como insumo básico para la ejecución de las tareas y actividades cotidianas.

Las unidades de información especializadas en el área comercial, como la Red de Bibliotecas de la Cámara de Comercio de Bogotá, el Centro de Documentación de la Superintendencia de Sociedades, las bibliotecas universitarias en apoyo a sus programas de pregrado, posgrados, especializaciones, maestrías y doctorados en áreas comerciales, entre otros. Se evidencia que no tienen un tesoro el cual les permita normalizar la terminología empleada para referirse a las actividades comerciales de Colombia, dificultando sus procesos de migración, interoperabilidad y reutilización de la información; a la par, que imposibilita la visibilidad y acceso a dicha información por parte de los usuarios; ya que instrumentos afines a la temática como lo son: el Tesoro multilingüe de la Unión Europea – EUROVOC, el Tesoro OIT, el Tesoro SECTUR del sector turismo (Argentina), el THESAURUS de la Biblioteca de la Empresa Familiar Carles Ferrer Salat y el directorio temático del INTAL, no son suficientes para apoyar y suministrar información en lo referente con las actividades comerciales colombianas por no tener en su estructura este alcance y profundidad en la temática, lo que conlleva al empleo de instrumentos con gran limitación terminológica como es el caso de las LEMB (Listas de Encabezamientos de Materia) o la construcción sin criterio y poco normalizada de palabras claves

locales acordes a cada tipo de Unidad de Información, hecho que dificulta la interoperabilidad y reutilización de los recursos de información.

Razones que dan vía libre a la viabilidad de realizar este trabajo investigativo como un aporte disciplinar en lo concerniente con estructuración, análisis y recuperación de información en lo referente a las actividades comerciales empleadas en Colombia.

Por eso me planteo el interrogante: ¿Existe coherencia y pertinencia en los términos empleados para indizar y recuperar los documentos que desarrollan temáticas o aspectos del ámbito comercial?

2. OBJETIVOS

2.1 OBJETIVO GENERAL

Construir un tesaurus especializado en el área comercial, a partir del marco normativo vigente en materia de tesaurus.

2.2 OBJETIVOS ESPECÍFICOS

- Identificar las fuentes documentales correctas para la construcción del tesaurus en materia de actividades comerciales empleadas en Colombia.
- Definir principios y criterios que rijan una adecuada construcción de tesaurus.
- Aplicar los parámetros técnicos y semánticos dados desde el marco normativo existente en lo referente con la construcción de Tesaurus.

3. JUSTIFICACIÓN

La investigación propuesta, buscó puntos de encuentro y contraste mediante la revisión de la teoría y los conceptos existentes sobre tesauros comerciales, y es por esta razón se evidencia una inexistencia de control y normalización terminológica en la información del sector comercial, hecho que consideramos causa grandes inconvenientes y obstáculos para la recuperación de la misma, razón por la cual se hace necesario la construcción de un Tesauro 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 este dominio específico, permitiéndonos de esta manera, efectuar intercambios de información entre las entidades del sector y el desarrollo terminológico para esta área de conocimiento.

Por todo lo anterior, hoy en día es más latente la necesidad de emplear lenguajes documentales en actividades relacionadas con la organización, clasificación, representación, marcaje y recuperación de información, especialmente en entornos web, lo anterior 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.

A partir de lo anteriormente expuesto, podemos ver implícita nuestra gran responsabilidad social y profesional, de generar potentes herramientas e instrumentos de gestión de información y conocimiento, que permitan generar valores agregados a nuestras actividades inherentes al procesamiento de información y que coayuden a 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.

Debido a la necesidad de analizar e indizar el conocimiento por parte de los profesionales de sistemas de información, bibliotecas especializadas en áreas interdisciplinarias, se detectó la necesidad de construir un tesoro en el área comercial, sobre las actividades comerciales empleadas en Colombia ajustado a la descripción, clasificación, indexación y búsqueda de los términos comerciales, con el fin de que los empresarios y los usuario puedan acceder y recuperar la información más fácilmente, coayudando a suplir de esta manera sus necesidades informativas.

Se evidencia la carencia de tesauros similares o afines a la temática de este tesoro, ya que solo se pudieron identificar de forma relacionada pero no cercana, ni afín, los siguientes tesauros:

- Tesoro multilingüe de la Unión Europea – EUROVOC, es un tesoro de la Unión Europea y su ámbito terminológico está orientado a labores parlamentarias de su país.
- Tesoro de la Organización Internacional del Trabajo (Tesoro OIT), el cual contiene una recopilación de 4,000 términos relacionados con el mundo del trabajo de las Naciones Unidas.
- Tesoro del Sector Turismo (Tesoro SECTUR - Argentina), con la elaboración del glosario de terminología turística sectorial de Argentina.
- El THESAURUS de la Biblioteca de la Empresa Familiar Carles Ferrer Salat, el cual persigue acotar y establecer las relaciones existentes entre ellos para reflejar la realidad del área de conocimiento de empresarios y emprendedores.
- El directorio temático del INTAL el cual comprende 11 temas afines al área comercial entre otros, pero no cuentan con un área específica de términos comerciales.

Ante el panorama anteriormente descrito, se ve la necesidad de innovar con la construcción de un tesoro en el área comercial, sobre las actividades comerciales empleadas en Colombia, apoyado en la Clasificación Industrial Internacional Uniforme de todas las actividades Económicas para Colombia (CIIU), emitido por la DIAN y las Cámaras de Comercio, el cual es tomado como referente para la validación de descripciones en las actividades económicas adaptadas para Colombia.

Con el tesoro sobre el área comercial, apoyado en las actividades comerciales empleadas en Colombia, se pretende homogenizar, normalizar y estructurar la terminología de esta área del conocimiento, bajo un lenguaje controlado que contribuya a la sociedad como una herramienta de consulta, y que establezca adecuadamente las relaciones existentes entre ellos, para que no sea solo un listado de significados terminológicos sueltos, aislados, sin análisis y sin estructura.

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.

4. DISEÑO METODOLÓGICO

4.1 TIPO DE INVESTIGACIÓN

Esta investigación está adscrita a la línea de investigación: información, conocimiento y sociedad; la sub línea Gestión de la Información y el Conocimiento, ya que aporta elementos para el estudio de la información y el conocimiento, considerándoles como un recurso estratégico para el avance social de los habitantes de una comunidad. A la par que desarrolla fundamentos teóricos y prácticos para el análisis, la interpretación y la clasificación de la información y el conocimiento registrado en los documentos en lo referente con el área comercial apoyado en las actividades comerciales empleadas en Colombia, permitiendo una fácil recuperación de la terminología a partir del vocabulario controlado.

La investigación se desarrolló a partir de un marco metodológico centrado en un enfoque de tipo cualitativo y con método deductivo, las técnicas que se emplearon para la recopilación de datos fueron las documentales, 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”².

Para la recolección de información se emplearon los fundamentos de la investigación documental como la consulta de fuentes primarias, identificación de especialistas en Tesauros y en el tema de Comercio, así como, las experiencias de investigación cualitativas y cuantitativas. Nuestra fuente primaria de

² 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>

investigación fue la Clasificación de Actividades Económicas – CIIU revisión 4 adaptada para Colombia, mediante la resolución No. 000139 de Noviembre 21 de 2012 y contamos con la asesoría de los siguientes especialistas, en materia de Tesoros el Señor Edwin Sossa, docente de la Universidad de la Salle, Claudia Rodríguez redactora de los códigos de Comercial, Tributario y Contable en Legis Editores y las Dras. Carolina Rangel y Susana Echeverri, abogadas especialistas en materia de Derecho Comercial, coordinadoras de publicaciones electrónicas en Legis Editores.

La búsqueda de los términos del tesoro se realizó bajo el método deductivo / analítico; este método se basa en el análisis del contenido de un conjunto de documentos elegidos y suele utilizarse en campos específicos, por la gran ventaja de ofrecer un cumulo de terminología con sentido real y en uso.

Según la AFNOR (Organización Nacional para la Estandarización) 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.
- Los términos indizados por acumulación.

La metodología para la construcción del tesoro y su desarrollo, siguió las recomendaciones establecidas por las normas:

1. La Norma ISO 25964 - el estándar internacional para tesoros y la interoperabilidad con otros vocabularios.
2. La Norma Técnica Colombiana NTC 1476:1998 – Guía para el establecimiento y desarrollo de tesoros monolingües de ICONTEC.
3. La norma Española UNE 50 – 106 – 90 Directrices para el establecimiento y desarrollo de tesoros monolingües.

4. La Norma ISO 704:2009: Terminology work -- Principles and methods
5. La Norma ISO / TC 37 - Terminología y otros recursos lingüísticos y de contenido.
6. La Norma ISO 860: Trabajo de Terminología - Armonización de conceptos y términos
7. La Norma ISO 29383: Políticas Terminología - Desarrollo e implementación.

FASES METODOLÓGICAS DEL TESAURO

La construcción del tesauro en el área comercial, sobre las actividades comerciales empleadas en Colombia, además de tener en cuenta las anteriores recomendaciones se incluye una revisión de políticas o facetas, para determinar la su estructura inicial, la recopilación, selección y relaciones de términos, cobertura temática y los límites del lenguaje a utilizar de acuerdo con las siguientes observaciones:

1° Estructura inicial: Se realizó una exhaustiva investigación y análisis de los documentos especializados sobre el tema bajo el contexto de los diferentes trabajos realizados sobre la elaboración de tesauros y sus experiencias en el mismo, teniendo en cuenta los diferentes usos terminológicos de empresarios y temas afines. Lo que permitió realizar una estructura inicial en un esquema para incorporar los términos, indicando los tipos de conexión, niveles de conexión y estructura jerárquica utilizada.

2° Recopilación de los términos: La recolección de los términos en la construcción del tesauro en el área comercial, se realizó bajo el método deductivo que parte de datos generales aceptados como válidos para llegar a una conclusión de tipo particular, con algunos componentes del método analítico que consiste en la separación de las partes de un todo para estudiarlas en forma individual. Con estos dos métodos se logró el análisis del contenido terminológico en la

construcción del tesoro por medio de la elaboración del listado de términos principales.

3° Selección de términos: La agrupación de las listas de los términos se realizó bajo una estructura facetada y apoyada con los especialistas, lo anterior con el fin de recoger sus puntos de vista y poder identificar los términos más importantes y relevantes. Como se mencionó en la justificación, son varias las herramientas terminológicas que se encuentran ya a disposición de los usuarios, sin embargo pocas las enfocadas directamente a las actividades comerciales empleadas en Colombia.

A continuación se relacionan algunas que tratan temas relacionados sin enfoque empresarial: el Tesoro multilingüe de la Unión Europea – EUROVOC³, es un tesoro de la Unión Europea y su ámbito terminológico está orientado a labores parlamentarias de su país, también está el de la Organización internacional del trabajo el tesoro de la OTI⁴, el directorio temático del INTAL⁵, el CIIU Rev. 4 A.C.6, el Tesoro SECTUR⁷, THESAURUS⁸ de la Biblioteca de la Empresa Familiar Carles Ferrer Salat, entre otros; con el fin de generar y/o definir una lista de términos elegidos y ordenados en forma alfabética y temática.

4° Relaciones entre términos: Esta es la parte donde a cada término se le asignaron las relaciones de equivalencia, las relaciones asociativas, las notas explicativas (si hubiese lugar) y las relaciones jerárquicas que se requirieran, bajo los lineamientos dados en el marco normativo anteriormente citado.

³ EUROVOC. Op. cit.,

⁴ EL TESAURO DE LA OTI. Op. cit.,

⁵EL DIRECTORIO TEMÁTICO DEL INTAL – BID . Op. cit.,

⁶ CIIU Rev. 4 A.C, Op. cit.,

⁷ SECTUR, Op. cit.,

⁸ THESAURUS. Op. cit.,

5° Elección de descriptores: Para elegir los términos descriptores que componen el Tesauro, se tuvieron en cuenta algunos elementos como:

- El término más usado comúnmente.
- La amplitud del término.
- Los grados de ambigüedad.
- La estructura morfológica y semántica del término.
- Las formas del término (sustantivo, sintagma nominal, singular, masculina y de escritura directa)
- El empleo de los plurales se reservó para objetos físicos, materiales, palpables y contabilizables.
- Se evitó el abuso de preposiciones y se excluyeron signos de puntuación, diacríticos, caracteres especiales y abreviaturas.

6° Revisión y edición: Antes de la edición se realizó una revisión por parte del especialista, con el fin de detectar posibles inconsistencias o errores, lo que dio paso a la realización de los posibles cambios y ajustes, con el fin de perfeccionar la estructura del Tesauro en el área comercial.

Además de colaborar con instituciones afines, este proyecto pretende (con el uso del tesauro), analizar e indizar adecuadamente la producción de información especializada en el área comercial de las diferentes unidades de información especializada en temas de comercio y que estén en armonía con la descripción de las actividades económicas del Código CIIU Rev. 4 A.C9, con el fin de contribuir a que la información procesada y/o registradas, pueda contar con una estructura coherente, organizada y de fácil acceso al momento de catalogar o recuperar su información.

⁹ CIIU Op. cit.

La estructura general del Tesauro en el área comercial, se implementó sobre el software libre TemaTres¹⁰, el cual sirvió de base para la presentación del producto final, también se empleó e implemento la planilla para el análisis semántico sugerida por la norma NISO Z39.19-2005 (R2010)¹¹, garantizando un efectivo acopio y descripción de cada uno de los términos tratados en el tesauro en el área comercial. Alternamente se construyó una taxonomía la cual permitió organizar la información de lo particular a lo general, permitiendo visualizar los términos generales y sus específicos, como insumo básico y esencial de la entrega final del producto.

RECURSOS PARA LA CONSTRUCCIÓN DEL TESAURO EN AREA COMERCIAL

Se tendrá en cuenta:

Glosarios, Diccionarios y Manuales

Los glosarios, manuales y diccionarios especializados que contengan términos referentes con las actividades comerciales realizadas en Colombia, ayudándonos a ubicar los significados de cada uno de los descriptores que conformaran la estructura general del tesauro, los cuales se encuentran en las Bibliotecas Universitarias y centros de información especializados de la ciudad.

Tesauros en línea afines al área comercial

Se identificaron cuatro tesauros y un directorio de terminología afín a la temática del Tesauro, los cuales fueron:

¹⁰ TEMA TRES Op. cit.,

¹¹ NISO Z39.19-2005 (R2010). Directrices para la construcción, el formato, y gestión de vocabularios controlados monolingüe. ESTÁNDAR PUBLICADO 05/13/2010 POR la Organización Nacional de Normas de Información. En línea:
<http://translate.google.com.co/translate?hl=es&sl=en&u=http://www.techstreet.com/products/1171385&prev=/search%3Fq%3DANSI/NISO%2BZ39.19%2B%25E2%2580%2593%2B2003%26biw%3D1344%26bih%3D846>

Tesaurus multilingüe de la Unión Europea – EUROVOC¹²

En temas de Vida política, Relaciones internacionales, Unión europea., Derecho, Vida económica, Intercambios económicos y comerciales, Asuntos financieros Asuntos sociales, Educación y comunicación, Ciencia, Empresa y competencia, Trabajo y empleo, Transportes, Medio ambiente, Agricultura, silvicultura y pesca, Sector agroalimentario, Producción, tecnología e investigación, Energía, Industria, Geografía, Organizaciones internacionales.

Organización internacional del trabajo. Tesauro OIT¹³,

En temas de relaciones internacionales, política social, protección social y seguridad social, desarrollo económico, derecho, derechos humanos, gobierno y política, ciencias sociales, cultura, humanidades y arte, educación y formación, desarrollo rural, agricultura, silvicultura y pesca, actividades económicas, comercio, transporte, financiamiento administración, trabajo y empleo, población, relaciones raciales y migración, salud y seguridad, ciencias ambientales, ciencias de la tierra, investigación y ciencia, Información y ciencias de la información.

EI SECTUR¹⁴ tesauro del sector turismo (Argentina), con la elaboración del glosario de terminología turística sectorial de Argentina,

EI THESAURUS¹⁵ de la Biblioteca de la Empresa Familiar Carles Ferrer Salat persigue acotar y establecer las relaciones existentes entre ellos para reflejar la realidad del área de conocimiento... de empresarios y emprendedores.

¹² EUROVOC. Op. cit.,

¹³ OIT. Op. cit.,

¹⁴ SECTUR. Op. cit.,

¹⁵ THESAURUS. Op. cit.,

El directorio temático del INTAL¹⁶ el cual comprende 11 temas afines al área comercial entre otros pero no cuentan con un área específica de términos comerciales. Entre otros.

Estas herramientas fueron empleadas como referente bibliográfico para verificar y validar la terminología rastreada a partir del método deductivo – analítico, seguido en la construcción del tesauro en el área comercial.

En términos generales el aporte al proceso fue poco ya que mucha de la terminología rastreada no se evidenciaba en estas fuentes por no ser una terminología muy enfocada o direccionada al contexto colombiano.

Herramienta para la Automatización del Tesauro en el área comercial

Para la construcción del tesauro en el área comercial, se requiere de una herramienta gestora que permita una mejor interacción entre la terminología recopilada y los usuarios, de forma tal que se adapte a las necesidades de los usuarios, que responda a los aportes requeridos por las normas y que sea usable y amigable. Desde esta perspectiva iniciamos un proceso de valoración en el mercado sobre los recursos disponibles y encontramos gran variedad de empresas con software que permiten la construcción de este tipo de herramientas, tanto comerciales como libres y que se ajustan a lo requerido desde el componente técnico y funcional.

Desde el panorama anteriormente descrito, se procedió con la evaluación de tres aplicativos para la construcción de tesauros que cuentan con las características requeridas por las normas, los cuales pueden ser objeto de análisis y comparación, estas tres herramientas fueron: TemaTres, MultiTes, Data Harmony (Synaptica), evaluación que fue realizada bajo los requisitos funcionales planteados a partir de los siguientes referentes:

¹⁶ INTAL. Op. cit.,

1. La Norma ISO 25964 - el estándar internacional para tesauros y la interoperabilidad con otros vocabularios.
2. La Norma Técnica Colombiana NTC 1476:1998 – Guía para el establecimiento y desarrollo de tesauros monolingües de ICONTEC.
3. La norma Española UNE 50 – 106 – 90 Directrices para el establecimiento y desarrollo de tesauros monolingües.
4. La Norma ISO 704:2009: Terminology work -- Principles and methods
5. La Norma ISO / TC 37 - Terminología y otros recursos lingüísticos y de contenido.
6. La Norma ISO 860: Trabajo de Terminología - Armonización de conceptos y términos
7. La Norma ISO 29383: Políticas Terminología - Desarrollo e implementación.

Para llevar a cabo dicha evaluación, se utilizó una serie de criterios específicos para este tipo de software: por un lado, los criterios relativos a la construcción y mantenimiento de tesauros (definición de términos y sus atributos, relaciones entre los términos, edición y mantenimiento del tesauro); por otro lado, criterios sobre las salidas que ofrecen estos paquetes de softwares (salida a pantalla, salida a impresora, consulta, navegabilidad, exportación e interoperabilidad). Los resultados de dicha evaluación se presentan teniendo en cuenta los siguientes criterios: (Ver anexo B del presente trabajo).

Datos comunes al software evaluado

El software para la gestión de tesauros, se puede dividir en dos grandes grupos: programas independientes de cualquier otro tipo de software o programas que sólo están disponibles como un módulo de un sistema de almacenamiento y recuperación de información. Aunque hay una serie de ventajas al utilizar un módulo de gestión de tesauros dentro de un sistema completo (integración entre

las operaciones de catalogación, indización y búsqueda), hay que tener en cuenta que dichos sistemas necesitan de muchos más recursos hardware (memoria principal, espacio de almacenamiento, velocidad de procesamiento, etc.) y software (sistemas gestores de bases de datos), que los sistemas de gestión de tesauros independientes. Estos últimos son más apropiados cuando no se necesita una interacción con el sistema de almacenamiento y recuperación de la información, permitiendo mayores niveles de interoperabilidad y una administración mucho más independiente y en tiempo real sobre el mismo, al no requerir de procesos complejos de exportación y control.

Se distinguen tres grandes bloques sobre los criterios para el análisis de estas herramientas, los cuales veremos con mayor detenimiento:

Bloque 1: Construcción y mantenimiento del tesoro

La construcción y el mantenimiento de un tesoro incluye una variedad de tareas y actividades, entre las que podemos destacar:

Los términos que constituyen el tesoro se seleccionan y almacenan en el mismo sistema. Es aconsejable guardar junto a ellos, una serie de informaciones relativas a los mismos (atributos) como, por ejemplo, el idioma, la fuente del término, las notas de alcance, definiciones, etc. En ocasiones deben integrarse términos procedentes de diferentes tesauros previamente existentes. Generalmente, este tipo de software está diseñado para soportar las relaciones básicas:

- Relaciones de equivalencia: USE/UF.
- Relaciones jerárquicas: BT/NT.
- Relaciones asociativas RT/RT.

También están diseñados para mantener la consistencia en el tesauro, asegurando que cuando se crea o se elimina una de estas relaciones entre dos términos, la relación recíproca también se crea o se anula.

Por último, el software debe ofrecer facilidades para el mantenimiento del tesauro, puesto que estos deben actualizarse regularmente con respecto a las relaciones y terminología debido a los cambios o mutación de las palabras en el tiempo o el contexto de uso, o por aspectos de usabilidad relacionados con los hábitos en la búsqueda y recuperación por parte de los usuarios.

Bloque 2: Salidas del tesauro

Dentro de este análisis, se tiene en cuenta:

Visualización en la pantalla del tesauro, salidas impresas y salida o volcado a un archivo. En cuanto a la visualización en pantalla del tesauro, o de una parte seleccionada previamente, todas las aplicaciones de gestión de tesauros permiten la visualización del mismo de forma alfabética y sistemática. Normalmente también ofrecen una visualización jerárquica y muchos de ellos incluso una permutada (KWIC/KWOC). Las salidas por la impresora suelen brindar las mismas alternativas que se dan para la visualización en pantalla. Por último, la mayoría de los aplicativos existentes para la gestión de tesauros, presentan la salida a un archivo, que se puede utilizar para su posterior edición por parte de un procesador de textos, o incluso para la integración de los datos en otro tesauro.

TemaTres: Permite implementar modelos distribuidos de gestión garantizando condiciones de consistencia e integridad de datos y relaciones entre términos y está desarrollada para ofrecer datos de trazabilidad y control de calidad en el contexto de un vocabulario controlado, con capacidades y funcionalidades de consulta, análisis y asistencia para la búsqueda y contextualización de términos, posibilita a su vez la representación de vocabularios a través de todos los estándares de metadatos disponibles en el campo de la gestión del conocimiento,

ofrece una interfaz de servicios web (web services) que permite y facilita su integración o articulación con otras plataformas de gestión o el desarrollo de servicios derivados basados en la explotación de vocabularios controlados.

En relación a los requerimientos técnicos para su instalación y funcionamiento se requiere aquello que la mayoría de los servicios de alojamiento web (hosting) ofrecen por defecto. En detalle:

- Un servidor web: mayoritariamente corre bajo Apache, pero puede funcionar bien en cualquier servidor web que soporte PHP, como el IIS (Internet Information Server) de las plataformas Windows.
- Una instalación de PHP en funcionamiento (versión 4.3.0 o posterior). PHP 5 está soportado.
- Un servidor de bases de datos (Mysql, Postgres, mssql, MS server, Oracle, etc.)

Dentro de sus características más destacables podemos mencionar:

- Escalabilidad: cuenta en la actualidad con más de 1.000.000 de términos
- Arquitectura orientada a servicios: API a través de Web services flexible y completa
- Compatibilidad: se puede exportar en una gran variedad de formatos: Skos-Core, Zthes, TopicMap, Dublin Core, MADS, BS8723-5, RSS, SiteMap, txt, SQ.

El aplicativo TemaTres ofrece una gran ayuda al usuario final, en la medida que presenta su documentación a través de una wiki la cual puede ser consultada en internet, su correspondiente manual y además ofrece una versión de prueba en línea. Dispone de dos módulos para la generación de reportes de editor que pueden ser utilizados como una herramienta para las rutinas de auditoría y control de calidad que se deseen implementar sobre el vocabulario controlado. Se incluye la posibilidad de emitir listados en formato *.CSV (archivo separado por comas,

para Excel, por ejemplo) en informes predefinidos (informes de términos libres, informes de términos repetidos, informes de términos con polijerarquía, informes de términos según estados) y a través del módulo de reportes avanzados se pueden establecer condiciones y filtros para la producción de informes de términos.

Adicionalmente posee una amplia gama de herramientas asociadas (Thesaurus WebPublisher, Tematres keywords distiller, WordPress WP-TemaTres, TemaTres & PKP Open Journal Systems, TemaTresView y Skoser Skoser).

Tema Tres es una herramienta que continuamente está desarrollando mejoras y nuevas funcionalidades, y en la actualidad se encuentran trabajando en un plugin para Omeka, una interfaz completamente jQuery y otra RDFa.

Bloque 3: Requerimientos de la herramienta TemaTres

Los requerimientos de TemaTres son los siguientes:

Un servidor web. La mayoría de los usuarios emplea Apache, pero TemaTres debe funcionar bien en cualquier servidor web que soporte PHP, como el IIS (Internet Information Server) de las plataformas Windows.

Una instalación de PHP en funcionamiento (versión 4.3.0 o posterior). PHP 5 está soportado.

Un servidor de bases de datos. Por defecto TemaTres utiliza MySQL, pero puede utilizar una variedad de motores de bases de datos (Postgres, Oracle MS server, MySQL, etc). Detalles acerca de los servidores de bases de datos soportados

Por todo lo anterior, se decidió emplear en la construcción del tesauro en el área comercial de este producto, el software libre “Tema Tres¹⁷”, ya que esta es una aplicación Web para la gestión de lenguajes documentales, su sitio oficial de descarga en: <http://www.vocabularyserver.com/comercio/> es de fácil uso, económicamente no genera costos de desarrollo, permite desarrollar estructuras de navegación Web y a su vez puede articular bibliotecas digitales con un gestor de contenidos.

4.2 DELIMITACIÓN Y ALCANCE

Para el desarrollo del presente trabajo de investigación se tomaron como guía:

1. La Norma ISO 25964 - el estándar internacional para tesauros y la interoperabilidad con otros vocabularios.
2. La Norma Técnica Colombiana NTC 1476:1998 – Guía para el establecimiento y desarrollo de tesauros monolingües de ICONTEC.
3. La norma Española UNE 50 – 106 – 90 Directrices para el establecimiento y desarrollo de tesauros monolingües.
4. La Norma ISO 704:2009: Terminology work -- Principles and methods
5. La Norma ISO / TC 37 - Terminología y otros recursos lingüísticos y de contenido.
6. La Norma ISO 860: Trabajo de Terminología - Armonización de conceptos y términos
7. La Norma ISO 29383: Políticas Terminología - Desarrollo e implementación.

Con el objeto de garantizar una práctica coherente en los criterios y técnicas empleadas en la construcción de tesauros, especialmente en la Construcción del

¹⁷ TEMA TRES Op. cit.,

Tesaurus en el área comercial, apoyado en las actividades comerciales empleadas en Colombia.

Se delimita con las siguientes pautas como: el análisis semántico, lingüístico, estandarizar e indexar el uso de términos a través de descriptores y no descriptores, entre los que establezcan una relación de equivalencia para que se incluyan en el mismo y que los no descriptores remitan al usuario al descriptor adecuado. Lo anterior, con el fin de garantizar una práctica del desarrollo del producto coherente basada en los principios de la norma aplicables en cualquier campo temático para la construcción del tesaurus en área comercial, sobre las actividades comerciales empleadas en Colombia.

4.3 POBLACIÓN OBJETO

Está compuesta por el conjunto de documentos que registran la terminología propia del dominio en el cual se desarrollará el tesaurus del área comercial. De allí se tomarán los trazos descriptivos (terminología) para la posterior fijación de relacionamiento semántico, administración mediante software de gestión de tesauros y generación de la versión beta (de prueba) que aquí se presenta.

5. MARCO TEÓRICO

5.1 LENGUAJES DOCUMENTALES

Después de consultar diferente 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 1. Definiciones de lenguajes documentales.

AUTOR	DEFINICIÓN
GIL URDICIÁIN, 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 ¹⁸ .

¹⁸ GIL URDICIÁIN, Blanca. Manual de lenguajes documentales. Madrid: Ediciones Trea, 2004. P.28

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 ¹⁹ .
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 ²⁰ .
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 ²¹ .
NAUMIS PEÑA, Catalina:	Los lenguajes documentales no se utilizarán únicamente para indizar y recuperar, sino que es un elemento organizativo de los contenidos documentales ²² .

Fuente: El propio autor

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:

¹⁹ RODRÍGUEZ DELGADO. Rafael. Teoría de sistemas y tesauros facetados. EN : Revista Española de Documentación Científica, vol. 1, No. 3. (dic, 2005) P. 213-224

²⁰ 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>

²¹ INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS Y CERTIFICACIÓN. Guía para el establecimiento y desarrollo de tesauros monolingües. Bogotá : ICONTEC, 1998. P.15 (NTC 1476)

²² NAUMIS PEÑA, Catalina. Los tesauros documentales y su aplicación en la información impresa, digital y multimedia. México. Universidad Autónoma de México. 2007. P.113

- 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.
- 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 ó 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”²³, por ende el conocimiento y la información no son ajenos a su organización mediante las descripción 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 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*²⁴, divide este

²³ CURRÁS PUENTES, Emilia. Ontologías, taxonomías y tesauros: manual de construcción y uso. Gijón, Ediciones trea. 2005. P. 26

²⁴ 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 2008]. Disponible en: <<http://dialnet.unirioja.es/servlet/extaut?codigo=267457>> ISSN 0210-4164>

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.

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”²⁵.

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

²⁵ Ibid., P. 63

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 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”²⁶, 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.

²⁶ Ibid., P. 66

Estas clasificaciones, inspiradas en la lógica y en los sistemas filosóficos del conocimiento, se fundan en el principio de precoordination y son de carácter enciclopédico”²⁷.

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, 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 - SCDD²⁸ y la clasificación de la Biblioteca del Congreso - LCC²⁹ usaban

²⁷ GIL. Óp. cit., P. 30

²⁸ MONTGOMERY COUNTRY MEMORIAL LIBRARY SYSTEM. Sistema de Clasificación Decimal Dewey. [termino de búsqueda: SCDD], [en línea - html]. [consulta: 25 de ago., de 2008]. Disponible en :<www.countylibrary.org/span3.htm>

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 clasificación universal y se utiliza posteriormente como método de clasificación en áreas parciales de conocimiento”³⁰.

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.”³¹.

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 -LEMB³², que son comúnmente utilizadas en las bibliotecas para el uso de términos descriptores del contenido del

²⁹ 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>>

³⁰ 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 2013]. Disponible en: <www.hipertexto.info/documentos/xfml.htm>

³¹ 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>>

³² 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>

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 lenguaje en un sistema excesivamente “encorsetado” y rígido que impide las combinaciones en la recuperación”³³.

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”³⁴, 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.

³³ CRUZ RUBIO. Óp. cit., P.2

³⁴ GIL URDICIÁN, Óp. cit., P. 26

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

La combinación del uso del lenguaje natural es “representada por un vocabulario ilimitado de palabras y frases utilizadas habitualmente en un campo temático”³⁵, 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.

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”³⁶.

³⁵ LANCASTER, Frederick Wilfrid. Control del vocabulario en la recuperación de información. 2 ed. Madrid: Universidad de Illinois, 1995. P. 177.

³⁵ 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.

³⁶ NAUMIS. Óp. cit., P.115

El primero es el lenguaje sintético ó también conocido como precoordinado, son “sistemas fijos de clasificación general de todo el conocimiento existente en una época”³⁷. 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 contralado ó postcoordinado, el cual corresponde a un proceso de selección de términos directamente de las fuentes 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

³⁷ NAUMIS. Óp. cit., P.117

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 2. Ventajas y desventajas del uso de los lenguajes documentales.

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 necesidades terminológicas reales de los usuarios, y estos no deben realizar búsquedas sobre un lenguaje desconocido e infuncional característicos de la precoordinación.	Al existir diferentes denominaciones entre un mismo concepto, pueden existir la posibilidad de establecer relaciones incorrectas entre los términos.

<p>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.</p> <p>Existe multidimensionalidad de las relaciones entre términos descriptores, lo que amplía el espectro de un dominio y facilita la recuperación de información.</p>	<p>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.</p>
<p>La evolución de la terminología es más específica de acuerdo del área del conocimiento ó dominio.</p>	
LENGUAJE SINTÉTICO	
VENTAJAS	DESVENTAJAS
<p>A través del control terminológico se reduce significativamente la ambigüedad semántica mediante el control de lo homógrafos, lo que facilita establecer plenamente la representación de un término dentro de un dominio.</p>	<p>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.</p>
<p>Mayor uniformidad y constancia en la denominación de conceptos.</p>	<p>Generalmente no responden al lenguaje natural controlado requerido por de los usuarios y por tanto dificulta efectuar búsquedas en los términos utilizados por los usuarios.</p>
	<p>No existe multidimensionalidad entre términos debido a su estructura no combinatoria entre temáticas.</p>

Fuente: El propio autor

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 ó 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 ó 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 indización y recuperación información en el marco del Bibliotecología”³⁸.

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.

³⁸ 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 3. Términos de indización.

Términos de indización		
Clase	Alcance	Ejemplo
Entes concretos	Cosas y sus partes físicas	Bien
	Materiales	Vehículo
Entes abstractos	Acciones y eventos	Repoblación
	Entidades abstractas y propiedades de cosas, materiales ó acciones	Regulación
	Disciplinas ó ciencias	Economía
	Unidades de medida	Tasa de Cambio
Entes individuales	Expresados por nombre propios.	Colombia

Fuente: El propio autor

5.1.9 Términos simples y compuestos.

Para la construcción del tesauro en el área comercial, los términos se clasificaron en simples y 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:

- Vehículo

- **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:

Declaración de Bienes

Fabricación de Vehículos

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

Ejemplo:

Actividad Económica

Actividad Foco

Económica 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 perdida de su significado.

5.1.10 Términos en singular y plural.

A continuación se ilustra la forma correcta de descripción y empleo de los términos según el establecimiento de reglas semánticas.

Tabla 4. Uso de singulares y plurales

Clase	ENCABEZAMIENTO DE MATERIAS	UNE 50-106-90	Ejemplo
Singular	Conceptos abstractos.	Conceptos abstractos.	Personalidad.
	Fenómenos de la naturaleza.		Lluvia.
	Propiedades, condiciones ó características.	Propiedades.	Temperatura.
	Procesos ó actividades ejercidas ó sufridas por los seres vivos.	Actividades y disciplinas.	Fotosíntesis.
	Sistemas de creencias ó doctrinas.	Sistemas de creencias.	Comunismo.
	Entes concretos no cuantificables.		Sal.
	Partes u órganos del cuerpo singulares ó 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.	Bienes
	Grupo de personas como grupos étnicos, nacionales, religiosos y nombres de profesionales.		Colombianos.
	Partes y órganos del cuerpo dobles ó 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.

Fuente: El propio autor

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 5. 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.	Actividad de seguros Actividad Aseguradora
Polisémicos	Son aquellos que cuentan con “multiplicidad de aceptación de una palabra” ³⁹ ó 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)).

Fuente: El propio autor

³⁹ 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 6. 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 ó frases sustantivadas. Las frases sustantivadas pertenecen a la categoría de términos compuestos y aparecen en dos formas:</p> <p>Frases adjetivas</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>Ejemplos:</p> <p>PÁJAROS MARINOS BALONMANO</p> <p>Frases preposicionales</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.</p> <p>Cuando el modificador Contiene un adjetivo, debe preferirse como término candidato el sustantivo del que deriva. 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 MARES TR PÁJAROS MARINOS</p>	<p>Bajo frases nominales la norma entiende los términos compuestos: frases preposicionales ó frases preposicionales</p>
Adjetivos	Debe evitarse, en lo posible, el uso aislado de adjetivos en un lenguaje	El empleo de adjetivos aislados debe ser evitado.

	<p>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 ó 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 "Actividad" hacia términos compuestos como "Actividad Económica", "Actividad Agrícola", etc.</p>	<p>Los adjetivos de frases adjetivas pueden dar lugar a reenvíos.</p> <p>Ejemplo: FANERÓGAMAS (véanse las diferentes plantas fanerógamas)</p>
Adverbio	En otros idiomas como el inglés,	Adverbios tales como muy ó

	<p>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 excepto cuando esta adquiera un significado especial dentro de una jerga.</p>	<p>grandemente no son admitidos, salvo si la forman parte de expresiones técnicas. Ejemplo: MUY ALTA FRECUENCIA</p>
Verbos	<p>Los verbos expresados como participativos ó en infinitivos no deben utilizarse como términos de indización. Las actividades deben presentarse como sustantivos ó formas verbales sustantivadas. Ejemplo: ADMINISTRACIÓN (no "administrar ")</p>	<p>El infinitivo de los verbos no debe ser empleados solo como término de indización. Las actividades deben ser sustantivadas. Ejemplo: CONTABILIZACIÓN (no contabilizar)</p>
Abreviaturas, siglas y acrónimos	<p>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 ó se desconoce la forma completa del nombre. Aun así, deben mantenerse los reenvíos recíprocos entre una y otra forma. BANCO INTERAMERICANA DE DESARROLLO UP BID</p>	<p>Es preferible utilizar la forma desarrollada. Las siglas y acrónimos deben evitarse, salvo si se han sustantivado (LÁSER, RADAR) ó si se conocen en cualquier idioma (UNESCO)</p>

Fuente: El
propio
autor

	BID USE BANCO INTERAMERICANO DE DESARROLLO	
--	--	--

5.1.13 Selección de términos.

Para que exista un criterio de selección de términos utilizados en la construcción del tesaurus en el área comercial, es preciso conocer las posturas expuestas por la Norma UNE 50-106-90 y la Guía Metodológica para la Elaboración de Tesaurus creada por el ICFES.

Tabla 7. 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 ó cuasi-sinónimos.
Transliteración (Castellanización)	Cuando se incorporan términos de idiomas con alfabetos no latinos, debe seguirse el procedimiento descrito en las normas pertinentes.	Se hace referencias cruzadas (UP), seleccionando la palabra con la ortografía castellana y remitiendo a ésta el término de la ortográfica extranjera.
Términos de jergas y argot	Cuando existe un concepto de reciente aparición originado en un grupo ó subcultura particular y no se halla un término alternativo de aceptación	Cuando un concepto es expresado mediante un término originado dentro de una subcultura, grupo social ó región en particular y no

	<p>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>	<p>existen otra alternativa comúnmente aceptada, este debe ser aceptado, como término de indización.</p>
Denominaciones comunes y comerciales	<p>Con frecuencia, un producto se conoce por su 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>Frecuentemente a un producto se le conoce a 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	<p>Denominaciones coloquiales y científicas. Cuando un mismo concepto se designa por un término coloquial y</p>	<p>Cuando un nombre popular y uno científico se refiere al mismo concepto, se debe preferir aquella</p>

científicos	<p>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>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 ó 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>
Nombres de lugares (Toponímicos)	<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 tesauo. Se deben establecer referencias entre las formas preferidas y no preferidas.</p>
Nombres propios de instituciones y personas	<p>Con frecuencia, estos nombres se excluyen de un tesauo.</p> <p>Si se incluyen, se debe seleccionar la</p>	<p>Generalmente existe tipo de términos que se excluyen de un tesauo. Cuando se decide incluir,</p>

	<p>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>a) Los nombres de instituciones locales ó nacionales, es decir, aquellas que llevan a cabo sus actividades y publican quizás revistas ó 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 ó 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>la forma del nombre debe seleccionarse con base a los principios establecidos en las reglas angloamericanas de catalogación.</p>
--	---	---

	<p><i>Ejemplo (en un tesauro castellano):</i></p> <p>ORGANIZACIÓN DE LAS NACIONES UNIDAS</p> <p>UP UNITED NATIONS ORGANIZATION</p> <p>UP ONU</p> <p>United Nations Organization</p> <p>USE ORGANIZACIÓN DE LAS NACIONES UNIDAS</p> <p>ONU</p> <p>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></p> <p>JUANA DE ARCO</p> <p>UP Jeanne d 'Arc</p> <p>Jeanne d'Arc USE JUANA DE ARCO</p>	
--	---	--

Fuente: El propio autor

5.1.14 Definición de tesoro.

En la tabla que aparece a continuación se presentan algunas definiciones del término Tesoro.

Tabla 8. Definiciones del término Tesoro.

AUTOR	DEFINICIÓN
ISO 2788 Directrices para el establecimiento y desarrollo de tesauros monolingües:	“Según su función: instrumento de control terminológico que traduce a un lenguaje sistémico ó 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” ⁴⁰ .
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 ó 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” ⁴¹ .

⁴⁰ ORGANIZACIÓN INTERNACIONAL DE NORMALIZACIÓN. Documentation - Guidelines for the establishment and development of monolingual thesauri: ISO, 1986, P. 30. (ISO 2788)

⁴¹ 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” ⁴² .
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 URDICIÁIN, 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 reciprocas entre términos relacionados” ⁴³ .

⁴² SLYPE, Georges Van. Los lenguajes de indización: concepción, construcción y utilización de los sistemas documentales. Madrid : Fundación Germán Sánchez Ruipérez, 1991. P. 21.

⁴³ 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 conocimiento⁴⁴.</p>
--	---

Fuente: El propio 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 del 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.

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 conocimiento⁴⁵.

⁴⁴ INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS Y CERTIFICACIÓN. Óp. cit., P.66

⁴⁵ 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 tesauros documentales y su aplicación en la información impresa, digital y multimedia. México. Universidad Autónoma de México. 2007. P 122.

El término fue usado inicialmente para referirse a diccionarios e índices “con la misma connotación de lugar donde se guarda lo más importante”⁴⁶. El primero que uso la palabra “tesauro” en el campo de la literatura fue “Florentino Brunetto Latini en su enciclopedia *Livre du Trésor*”⁴⁷, en el siglo XIII. En la lengua española el primer diccionario fue creado por “Sebastián de Covarrubias reimpresso en 1672 por Gabriel de León”⁴⁸, sin embargo quien popularizo el término “tesauro” 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 Phrases Classified and Arranged so as to Facilitate the Expression of Ideas and Assist in Literary Composition"*⁴⁹, publicado en 1852.

Roget intentaba presentar la distinción existente entre un diccionario de un tesauro, debido a que el primero cuenta con un orden alfabético de palabras sin 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”⁵⁰, 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.

⁴⁶ *Ibíd.*, P.122

⁴⁷*Ibíd.*, P.123

⁴⁸ *Ibíd.*

⁴⁹ 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/>>

⁵⁰ GIL URDICIÁN. Blanca. Orígenes y evolución de los tesauros en España. *EN* : Revista General de Información y Documentación. Vol. 8, No. 1 (1998). P. 63 – 110.

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”⁵¹, 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 *ThesauroFacet: a thesaurus and faceted classification for engineering and related subjects* ó *Tesaurofacetas*, donde presenta “las ventajas de indizar y clasificar con una correspondencia entre facetas y descriptores”⁵², y Aitchison sustento que su tesauro “adoptaría una 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 tesauro, con las relaciones TE (término específico), TG (término genérico), TR (término relacionado), y el reenvió por un código de tres caracteres a la parte facetada, y la clasificación por materias, en que los términos son reagrupados en facetas ó funciones fundamentales”⁵³, lo que podría sustentarse como localización del término en un contexto alfabético y sistémico ó estructural de un dominio.

⁵¹ 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>>

⁵² CHAUMIER, Jacques. Análisis y lenguajes documentales: el tratamiento lingüístico de la información documental. Paris: Mitre. 1986. 12 p. 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 124.

⁵³ LÓPEZ, Miguel Ángel. Evolución histórica y tendencias observables en los tesauros. Departamento de Biblioteconomía y Documentación. Madrid. Universidad Carlos III de Madrid. 2006. P 363.

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 tesoro administrador por computadores que respondían a “modelización matemática”⁵⁴, 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 ingles) publica la norma 2788:1986 - Guidelines for the establishment and development of 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 tesoros monolingües⁵⁵. 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 tesoros 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,

⁵⁴ Ibid., P.365

⁵⁵ ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. UNE Directrices para el establecimiento y desarrollo de tesoros monolingües. Madrid: AENOR, 1990, P.40. (50-106)

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 –HTML⁵⁶ y Extensive Markup Language - XML⁵⁷ 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.

5.1.16 Funciones y uso del tesoro.

Según la autora GIL (2004), Los tesauros cumplen las siguientes funciones⁵⁸:

- 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

⁵⁶ Lenguaje de Marcas de Hipertexto para la construcción de páginas Web.

⁵⁷ Lenguaje de Marcas extensible), es un metalenguaje extensible de etiquetas desarrollado por el Word Wide Web Consortium (W3C).

⁵⁸ GIL URDICIÁN, Blanca. Manual de lenguajes documentales. Madrid (España) : Ediciones Trea , 2004.

podrían ser utilizados por los usuarios en las consultas, por lo tanto, está concebido en función de las necesidades de los usuarios.

- El tesauro 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 tesauro 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

Para poder lograr un entendimiento claro del Tesauro y sus componentes es necesario efectuar una conceptualización sobre sus componentes fundamentales, los cuales adquieren las categorías de descripción y de no descriptor. A continuación efectuaremos un análisis conceptual de ambos términos desde las perspectivas suministradas por varios autores.

- **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 ó 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”⁵⁹.

Catalina Naumis presenta varias consideración asertivas en referencia a las cualidades del término descriptor, las cuales se mencionan a continuación: ”el contenido ó 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”⁶⁰.

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.
- Cuando un descriptor puede ser expresado por varios sinónimos debe escogerse un único término como descriptor, que será aquel que desde el

⁵⁹ NAUMI. Óp. cit., P.134

⁶⁰ ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. Óp. cit., P. 5

punto de vista del idioma en que se está elaborando el tesoro responda a las siguientes características: propiedad, concisión y uso”⁶¹

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 ó 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”⁶².

a. Presentación de términos no descriptores:

⁶¹ INSTITUTO COLOMBIANO PARA EL FOMENTO DE LA EDUCACIÓN SUPERIOR. Oc. cit., P.37 – 38

⁶² ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. Óp. cit., P.5

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 ó las abreviaturas ó acrónimos, que siempre deban escribirse en mayúsculas”⁶³.

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 9. Clasificación de tesauros según su tamaño.

CATEGORÍA	DEFINICIÓN	EJEMPLO
MACROTESAUROS	<p>“Crear una especie de superestructuras genérica de términos que incluiría a un grupo de tesauros, u otros tipos de vocabularios de diversos campos temáticos”⁶⁴.</p> <p>“Macrotesauro es aquel que contiene diferentes temas relacionados entre sí, pero</p>	<p>Macrotesauro de la OCDE</p> <p>URL: http://bibliotecavirtual.clacso.org.ar/ar/oecd-macroth/es/</p> <p>(Consultado en: octubre 05 de 2015).</p>

⁶³ INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS Y CERTIFICACIÓN. Óp. cit., P.5

⁶⁴ LANCASTER. Op. cit., P. 215

	dispuestos en apartados diferentes, como si se tratase de un conglomerado de tesauros más pequeños” ⁶⁵ .	
TESAUROS	Dentro de la definición recolectada en las diferentes fuentes bibliográficas, no indican cuando puede considerarse tesoro según la cantidad de términos que lo componen.	Tesoro de redes de ordenadores URL: http://bartoc.org/en/node/165 (Consultado en: octubre 05 de 2015).
TESAUROS	“Tesoro no tiene definición universalmente aceptada. Con frecuencia se utiliza para hacer referencia a un tesoro pequeño y especializado. Puede definirse como un vocabulario especializado, que es compatible con algún tesoro más amplio y que está totalmente integrado en su estructura jerárquica” ⁶⁶ .	Tesoro de Huridocs URL: https://www.huridocs.org/wp-content/uploads/2010/07/mict_hespfinalpdf.pdf (Consultado en: octubre 05 de 2015).

Fuente: El autor.

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

⁶⁵ CURRAS. Óp. cit., P. 35

⁶⁶ LANCASTER. Op. cit., P. 212

Tabla 10. Clasificación de tesauros según su idioma.

Fuente: El propio autor.

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 Tesauro ISOC de Psicología URL: http://thes.cindoc.csic.es/index_PSICO_esp.php (Consultado en: octubre 05 de 2015).
Multilingües	Está compuesto por términos de diferentes idiomas y están orientados por la ISO 5964: 1985.	Ejemplo: Inglés, Francés, Alemán, Italiano, Español Astronomy Thesaurus URL: http://www.thesaurus.com/browse/astronomy (Consultado en: octubre 05 de 2015).

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.

Desde la perspectiva de Blanca Gil Urdiciain efectuaremos un análisis sobre las relaciones y la simbología empleada en los Tesoros, aspectos que desarrollaremos a continuación: ⁶⁷

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 – Término General) y (TE – Término Específico) en el idioma de Español, (BT – Broad Term o término más amplio) y (NT - Narrow Term o término más específico) en el idioma

⁶⁷ GIL URDICIAIN, Blanca. Manual de lenguajes documentales. Madrid (España) : Ediciones Trea , 2004.

Ingles respectivamente. Su objetivo es expresar su reciprocidad entre los términos genéricos y específicos como se presenta a continuación:

Tabla 11. Relaciones jerárquicas.

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.

Fuente: El propio autor.

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”⁶⁸.

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

⁶⁸ NAUMIS. Óp. cit., P. 195.

Tabla 12. Nota de alcance.

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

Fuente: El propio autor.

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 ó nota histórica.

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 tesauro, ya que este recurso ofrece

términos alternativos que podrían ser utilizados para la indización ó para la recuperación. Esta relación es reciproca”⁶⁹.

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

Tabla 13. 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 ó término específico.

Fuente: El propio autor.

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.

⁶⁹ INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS Y CERTIFICACIÓN. Óp. cit., P.40

Tabla 14. Relaciones de equivalencia.

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

Fuente: El propio 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 15. Relaciones y símbolos del tesauro.

	TIPO DE RELACIÓN	CONCEPTO	CARACTERÍSTICA	SÍMBOLOS	EJEMPLO
RELACIONES SEMÁNTICAS	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 tesauro.	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 tesauro 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 propio autor.

6. CONSTRUCCIÓN DEL TESAURO ESPECIALIZADO EN EL ÁREA COMERCIAL

6.1 METODOLOGÍA PARA LA CONSTRUCCIÓN DE UN TESAURO ESPECIALIZADO EN EL ÁREA COMERCIAL

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

www.vocabularyserver.com/comercio/index.php?letra=A

Inicio | Menú | Sobre... | Búsqueda avanzada

Buscar

TESAURO ESPECIALIZADO EN EL ÁREA COMERCIAL

A B C D E F G H I J L M N O P R S T U V Z

Inicio ▶ A: 144 términos

1. ACABADO DE EDIFICIOS
2. ACABADO DE OBRAS DE INGENIERÍA CIVIL
3. ACABADO DE TEXTILES
4. ACTIVIDAD ACUÍCOLA
5. actividad agraria **USE** ACTIVIDAD AGRÍCOLA
6. ACTIVIDAD AGRÍCOLA
7. ACTIVIDAD AGROINDUSTRIAL
8. ACTIVIDAD AGROPECUARIA
9. ACTIVIDAD AMBIENTAL
10. ACTIVIDAD ARTÍSTICA
11. ACTIVIDAD ASEGURADORA
12. ACTIVIDAD AUTOMOTRIZ
13. actividad bancaria **USE** ACTIVIDAD BURSÁTIL
14. ACTIVIDAD BURSÁTIL
15. ACTIVIDAD CIENTÍFICA
16. ACTIVIDAD COMERCIAL
17. ACTIVIDAD CONSTRUCTORA
18. ACTIVIDAD CONTABLE
19. ACTIVIDAD COSMÉTICA
20. ACTIVIDAD CULTURAL
21. ACTIVIDAD DE APOYO DIAGNÓSTICO
22. ACTIVIDAD DE APOYO TERAPÉUTICO
23. ACTIVIDAD DE ARQUITECTURA
24. ACTIVIDAD DE ARRENDAMIENTO
25. ACTIVIDAD DE ASOCIACIÓN
26. ACTIVIDAD DE CAZA
27. ACTIVIDAD DE COMERCIO
28. ACTIVIDAD DE CONSULTORÍA
29. ACTIVIDAD DE EDICIÓN
30. ACTIVIDAD DE EDUCACIÓN
31. ACTIVIDAD DE EMPAQUE
32. ACTIVIDAD DE ENVASE
33. ACTIVIDAD DE FOTOGRAFÍA
34. ACTIVIDAD DE INFORMÁTICA
35. ACTIVIDAD DE INGENIERÍA
36. ACTIVIDAD DE INVESTIGACIÓN

Tomada de: <http://www.vocabularyserver.com/comercio/index.php>. Consultada en octubre 05 de 2015.

6.1.1 Profundidad y extensión del Tesauro especializado en el área comercial.

El Tesauro fue construido a partir de la terminología empleada en cada sector o área de actividades comerciales empleadas en Colombia, su profundidad está limitada a la estructura y profundidad de la misma; sin embargo, este Tesauro fue diseñado con la finalidad de ser empleado por otras entidades que cumplen una labor semejante dentro del 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 tesauros.

Teniendo en cuenta que el Tesauro está desarrollado para su aplicación en el Sector Comercio, gran porcentaje de los términos recolectados se efectuó directamente sobre los documentos técnicos que generan las entidades del sector, tomando como referente Legis Editores y la Cámara de Comercio de Bogotá, 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 Tesauro, integrarlo y/o reutilizarlo con otros Centro Documentales que estén interesados en esta temática.

Para la construcción del corpus terminológico del Tesauro, 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 Tesauro inicialmente en áreas generales del conocimiento propuestas por los sistemas precoordinados como el Sistema de 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

El proceso de análisis es considerado una de las actividades más importantes dentro del proceso de construcción de Tesauros, ya que en esta actividad se centra y evidencia la destreza del analista para detectar y obtener los términos que harán parte integrante del Tesauro especializado en el área comercial; por consiguiente, este proceso requiere una persona con alto conocimiento no solo de las técnicas en materia de Tesauros, sino también un entendimiento y comprensión global del dominio o tema a tratar.

A continuación se describirán los pasos o aspectos a tener en cuenta al momento de realizar la captación y recopilación del acervo terminológico.

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 – 2003⁷⁰, 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 Tesauro en Comercio”*.

6.2.2 Selección y recolección de términos – corpus.

Por tratarse de un Tesauro 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.

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.

⁷⁰ 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.

Los términos extraídos se registraron en el *“Formato de Extracción y Actualización de Términos para el Tesauro en el área comercial”*, 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 Comercio.
- 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 pertenecientes al lenguaje natural, permitiendo un desarrollo de la estructura del Tesauro de lo general a lo específico.

Las facetas construidas para este Tesauro especializado en el área comercial fueron 9, categorizadas de la siguiente forma:

1. CLASES DE ACTIVIDAD ECONÓMICA
2. CLASIFICACIÓN DE LA ACTIVIDAD ECONÓMICA
3. CONTROL Y VIGILANCIA DE LA ACTIVIDAD ECONÓMICA

4. DEFINICIÓN DE ACTIVIDAD ECONÓMICA
5. EJERCICIO DE LA ACTIVIDAD ECONÓMICA
6. ELEMENTOS DE LA ACTIVIDAD ECONÓMICA
7. INFORMACIÓN DE LA ACTIVIDAD ECONÓMICA
8. INICIO DE LA ACTIVIDAD ECONÓMICA
9. MODALIDADES DE ACTIVIDAD ECONÓMICA

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 Excel⁷¹, esto con el propósito de agilizar el tratamiento de cada uno de los términos y proveer de un esquema taxonómico del Tesauro que facilitara la visualización de los términos por parte del grupo interdisciplinar que cooperó 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 tesauro.

Las fuentes técnicas consultadas fueron:

- Diccionarios especializados.
- Vocabulario controlado especializado.
- Sitios web especializados en el tema.

⁷¹ Excel es un software de Microsoft Office que le permite crear tablas, calcular y analizar datos.

- Repertorio legislativo sobre el tema.
- Documentos técnicos emitidos por organismos en la materia.
- Libros especializados en Comercio en el contexto latinoamericano.

Para acatar la metodología propuesta de validación, se conformo un grupo interdisciplinar de profesionales que cuentan con elementos teóricos y prácticos sobre el tema a desarrollar.

En lo que respecta a la validación de términos en forma y composición, el Director de trabajo de grado contribuyo 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 especializado en el área comercial, en relación a la cantidad de términos y establecimiento de relaciones:

- 988 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 2053 términos, debido a que estas unidades lingüísticas no correspondían temática al objeto que motivo la construcción del Tesauro especializado en el área comercial, 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 el caso, lo que convoca a esta investigación, adoptó los criterios preestablecidos que describiremos a continuación, los cuales deben ser tenidos en cuenta para la estructuración, mantenimiento y actualización de un tesoro homogeneizado⁷².

- **Criterio de Transmutabilidad**

Este criterio reconoce la adaptación, reinvención 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.

⁷² Aportaciones suministradas por el profesor Edwin Sossa.

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 Tesoros, es recomendable que los términos sean contruidos 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 siguiente⁷³:

- 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.

⁷³ Aportaciones suministradas por el profesor Edwin Sossa.

- **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 tesauro 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 Tesauro en el área comercial.

Para garantizar uniformidad en la construcción del Tesauro especializado en el área comercial, se conformaron las siguientes reglas semánticas, tomando como referentes aspectos normativos⁷⁴:

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.

⁷⁴ UNE 50-106-90. Directrices para el establecimiento y desarrollo de tesauros monolingües. [Consulta: 07 de ene., de 2014]. Disponible en: http://elmemorioso.cl/une50106_90.pdf

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 ó 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 ó 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 ó 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 ESPECIALIZADO EN EL AREA COMERCIAL

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 especializado en el área comercial, constituye en un elemento que facilita la participación, el intercambio de información, acceso y desarrollo terminológico en temas relacionados con Comercio, 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 especializado en el área comercial, deban estar ligados a esta corriente de libre acceso.

La licencia GPL (Licencia Pública General)⁷⁵, es un elemento que brinda la posibilidad de hacer uso del software para la construcción del Tesoro especializado en el área comercial sin incurrir en violaciones de derechos de autor.

Para éste producto 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 tesauros, el cual tiene por función de efectuar procesos de ingreso, actualización, consulta y eliminación de datos

⁷⁵ 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.

(términos), siendo éste el que presenta la interfaz y el contacto directo con el administrador y los usuarios.

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)⁷⁶ y en RDF (Resource Description Framework)⁷⁷.
- 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 Tesauro.
- Desarrollo en un ambiente PHP⁷⁸ que es un lenguaje GNU⁷⁹ lo que permite que opere eficientemente sobre bases de datos relacionales como MySql, Appserver ya sea en sistemas operativos Windows ó Linux.

⁷⁶ XML: Un lenguaje de marcado extensible que puede usarse para almacenar datos en un formato estructurado, basado en texto y definido por el usuario.

⁷⁷ 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.

⁷⁸ 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.

⁷⁹ GNU: Proyecto creado en 1984 con el fin de desarrollar un sistema operativo tipo UNIX según la filosofía del "software libre".

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 MADS⁸⁰, XML, SKOS-Core⁸¹, XTM⁸², DC⁸³.

⁸⁰ MADS: Esquema de metadatos de descripción y autoridad basado en XML.

- Los términos cuentan con hipervínculos que permiten navegar entre el Tesauro.
- 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 Tesauro, de acuerdo a la necesidad del administrador del tesauro especializado en el área comercial.
- 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.

Tema Tres

Este es un software libre que al momento de obtenerlo puede ser modificado, estudiado y su redistribución en libre, está disponible de forma gratuita en internet. TEMATRES es un software que permite implementar modelos distribuidos de gestión que garantizan las condiciones de consistencia e integridad de datos y relaciones entre términos, y está desarrollado para ofrecer datos de trazabilidad y control de calidad en el contexto de un vocabulario controlado. Además este se encuentra orientado al desarrollo de tesauros jerárquicos que se articulan con gestores de contenido, bibliotecas digitales o con una biblioteca tradicional.

⁸¹ SKOS-Core: Web semántica avanzada.

⁸² XTM: Es la fusión entre XML y topic maps, como medio de intercambio de metadatos.

⁸³ DC: Dublin Core es una organización abierta y libre para el desarrollo de estándares de metadatos interoperables en línea.

Ofrece una interfaz de servicios web (web services) que permite y facilita su integración o articulación con otras plataformas de gestión o el desarrollo de servicios derivados basados en la explotación de vocabularios controlados.

Funciones y Características:

Este maneja tres tipos de relaciones entre los términos:

- Relaciones jerárquicas
- Relaciones asociativas
- Relaciones equivalencia

Además es ilimitada la cantidad de términos, de niveles jerárquicos y de términos no preferentes. Cuenta con dos tipos de navegación:

Navegación Sistemática: La cual permite buscar un término por orden de jerarquía, siempre comenzando con el término más amplio hasta el término más específico, permitiendo navegar por toda la jerarquía.

Navegación Alfabética: Es otra manera de buscar un término, donde los descriptores y no descriptores están agrupados en una sola secuencia alfabética, donde están acompañados se los tipos de relaciones.

El Tesauro especializado en el área comercial está disponible para su consulta en la siguiente dirección: <http://www.vocabularyserver.com/comercio/> , cuenta con una estructura facetada compuesta 780 términos descriptores, 730 relaciones entre términos, 123 términos equivalentes y 12 notas de alcance.

www.vocabularyserver.com/comercio/index.php?tema=18/actividad-economica

Inicio | Menú | Sobre... | Búsqueda avanzada Buscar

TESAURO ESPECIALIZADO EN EL ÁREA COMERCIAL

ACTIVIDAD ECONÓMICA

Inicio ► ACTIVIDAD ECONÓMICA

- ✖ UP actividades comerciales
- ✖ UP actividades económicas

Opciones

ACTIVIDAD ECONÓMICA

- ✖ TE1 CLASES DE ACTIVIDAD ECONÓMICA ►
- ✖ TE1 CLASIFICACIÓN DE LA ACTIVIDAD ECONÓMICA ►
- ✖ TE1 CONTROL Y VIGILANCIA DE LA ACTIVIDAD ECONÓMICA
- ✖ TE1 DEFINICIÓN DE ACTIVIDAD ECONÓMICA
- ✖ TE1 EJERCICIO DE LA ACTIVIDAD ECONÓMICA ►
- ✖ TE1 ELEMENTOS DE LA ACTIVIDAD ECONÓMICA ►
- ✖ TE1 INFORMACIÓN DE LA ACTIVIDAD ECONÓMICA
- ✖ TE1 INICIO DE LA ACTIVIDAD ECONÓMICA
- ✖ TE1 MODALIDADES DE ACTIVIDAD ECONÓMICA ►

Fecha de creación: 15-Jun-2015
Término aceptado: 15-Jun-2015

BS8723-5 DC MADS SKOS-Core VDEX XTM Zthes JSON JSON-LD

URI: <http://www.vocabularyserver.com/comercio/>
API
Autor: Henry Castro Roa
español ▼

A B C D E F G H I J L M N O
P R S T U V Z

Tomada de: <http://www.vocabularyserver.com/comercio/index.php>. Consultada en octubre 05 de 2015.

Cuenta con impresión completa del tesauro especializado en el área comercial en XML: Zthes y también la impresión puede generar un archivo RDFS KOS CORE, que permite describir e intercambiar metadatos, y está construido en reglas de: recurso, propiedad y sentencia.

TEMATRES cuenta con tres tipos de notas:

Notas de alcance de los calificadores y detallaran el significado, y donde destaca que interpretaciones, restricciones, categorías y subcategorías son permitidas en cada calificador.

Notas de Históricas

Notas bibliográficas

Permite una búsqueda por término, permitiendo ingresar una parte o todo el texto del término deseado.

Tematres genera tres tipos de salidas:

Dublín Core: Es un sistema de 15 definiciones semánticas descriptivas que pretenden transmitir un significado semántico a las mismas. Estas definiciones Son opcionales, Se pueden repetir y Pueden aparecer en cualquier orden.

SKOS-Core: Es un esquema RDF para la representación de tesauros y sistemas similares de organización de conocimiento. Su objetivo es proporcionar un modelo para la migración de sistemas de organización de conocimiento al entorno de la Web Semántica.

Zthes: Este hace una descripción de un modelo abstracto para la representación y búsqueda de tesauros, de acuerdo con los con los términos que están definidos en la ISO 2788, en la cual especifica que los documentos se pueden definir según su función y según su estructura.

En cuanto a la administración de usuarios y editores Tematres cuenta con un módulo de autenticación de usuarios, donde estos deben loguearse con la dirección de correo electrónico y un clave, donde presenta dos tipos de acceso: 1.

Como usuarios para la búsqueda de información y 2. Como editores que sería para la carga del tesoro.

En cuanto a tecnología Tematres está desarrollado en PHP que es un lenguaje de programación cuya creación y desarrollo ha estado ligada siempre al producto GNU, opera sobre bases de datos Mysql. Requiere por lo tanto de un servidor HTTP con cobertura PHP y un servidor de bases de datos Mysql. Puede funcionar en cualquier tipo de plataforma (Linux, Windows). Si desea instalar la aplicación en un entorno Windows, se puede optar por instalar el paquete EasyPHP8, que trae ya integrados el servidor HTTP Apache9 y el servidor de bases de datos MySq.

Otras funciones son:

Utilidades de gestión y configuración

Auditoria de términos

Buscador de términos repetidos.

Buscador de términos libres

Cuenta con una interfaz multilingüe

Del análisis de los programas informáticos para la gestión de tesauros se concluye que:

1. En cuanto a las capacidades de almacenamientos de términos, se ha comprobado que el número máximo de términos que se pueden introducir por lo general no tiene limitaciones (salvo el espacio de almacenamiento disponible). En lo que se refiere a la longitud máxima de cada término, en casi todos los programas suele ser más que suficiente, si bien en algunos casos está limitada.

2. En lo que se refiere a la longitud de las notas se ha observado que todos cuentan con un espacio suficiente.
3. El control de la consistencia es un tema que por lo general suele estar bien resuelto, lo que ayuda a mantener la integridad y coherencia de los datos. No obstante, algunos de los programas presentan algunas como la formación de bucles en las relaciones jerárquicas.
4. Los sistemas trabajan con el servidor z39.50, protocolo que permite la migración e interoperabilidad de los datos entre diversos sistemas.
5. La mayoría de los programas evaluados son capaces de preparar ficheros para la exportación de datos a otros sistemas. En todos los casos, permiten la posibilidad de enviar a la impresora la visualización en pantalla que se está realizando.
6. TemaTres cuenta con información disponible para su instalación y administración, información que puede ser consultada en la siguiente dirección: <http://r020.com.ar/tematres/manual/>, adicionalmente el creador del programa, Diego Ferreira ofrece soporte sin ningún tipo costo para la instituciones que deseen implementarlo en la conformación de sus Tesauros.
7. Se recomienda implementar TemaTres ya que es un software Free e interoperable con diversas plataformas, lo que le daría la posibilidad a las diversas instituciones de poder migrar e interconectar el Tesauro sin tener necesariamente que adquirir una licencia, de igual forma al trabajar en ambiente WEB, los cambios efectuados sobre el Tesauro especializado en el área comercial, podrán ser vistos por las instituciones en tiempo real.

8. El Software TemaTres, por ser un software libre y open source tiene todos los beneficios que esto implica, como ser, es gratuito, se puede adaptarlo a las necesidades específicas modificando el código fuente e intercambiar opiniones con el autor (el cual está siempre dispuesto a colaborar) y otros usuarios para perfeccionar el funcionamiento de una herramienta tan económica y de tanta utilidad para las bibliotecas especializadas y centros de documentación.

7 CONCLUSIONES

El principal problema con la terminología de esta área del conocimiento consiste en la asignación, uso y significado de los mismos términos, ya que se obvia la complejidad que rodea a las unidades terminológicas empleadas, debido a que el uso va más allá de lo normalizado, va siempre por delante de lo normalizado y debido a esto, en el momento en que se normaliza, un término ya está obsoleto, porque el uso siempre va por delante de la normalización.

Otro principal problema se centra en la traducción inexacta de los términos, ya que gran parte de este corpus terminológico nace de términos extranjeros, lo que ocasiona problemas de adopción de los mismo según el contexto donde se van a emplear.

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 especializado en el área comercial, 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 tesoro 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 tesoro. Segundo, mediante la descripción, de manera general, del procedimiento básico para la elaboración de un tesoro 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.

8 RECOMENDACIONES

Para los profesionales en el área de la bibliotecología, que se interesan en la creación y el desarrollo de vocabularios controlados como los tesauros, se les recomienda evitar el uso de extranjerismos terminológicos ya que estos pueden dejar fuera de contexto donde se van a emplear, debido a la traducción inexacta de los términos.

Los aportes de innovación de nuevos tesauros, han de caracterizarse en la descripción metodológica de construcción y el diseño de la utilización de elementos que brinden nuevas soluciones, a través de la utilización de los gestores en línea adaptados al entorno tecnológico. Ya que constituyen una esencial pieza como medio de presentación y organización de los términos en la web semántica, permitiendo un entorno compartido y haciendo posible una interoperabilidad con otros vocabularios controlados.

BIBLIOGRAFÍA

CIIU Rev. 4 A.C. Clasificación Industrial Internacional Uniforme de todas las actividades económicas, adoptada y actualizada por el DANE para Colombia. [Consulta: 07 de ene., de 2014]. Disponible en: http://www.dane.gov.co/files/nomenclaturas/CIIU_Rev4ac.pdf

CURRÁS PUENTE, Emilia. Ontologías, taxonomía y tesauros : manual de construcción y uso. Gijón(Asturias) : Ediciones Trea,S.L. , 2005.

DANE - Departamento Administrativo Nacional de Estadística. [Consulta: 07 de ene., de 2014]. Disponible en: <http://www.dane.gov.co>

DEL VALLE GASTAMINZA, Félix. Lenguajes documentales: los tesauros. En: <http://www.ucm.es/info/multidoc/prof/fvalle/tesauro.htm>

EL DIRECTORIO TEMÁTICO DEL INTAL – BID – (Banco Internacional de Desarrollo), España. [Consulta: 07 de ene., de 2014]. Disponible en: <http://www10.iadb.org/intal/intalcdi/Tesauro.aspx?lang=es>

EL TESAURO DE LA OIT. España. [Consulta: 07 de ene., de 2014]. Disponible en: <http://www.ilo.org/public/spanish/support/lib/tools/aboutthes.htm>

ESPINOSA RICARDO, Lucy. Lenguajes de análisis de recuperación de Información. Armenia : Universidad del Quindío , 1989.

EUROVOC. Tesauro multilingüe y multidisciplinario de la Unión Europea, [consulta: 07 de ene., de 2014]. Disponible en: http://eurovoc.europa.eu/drupal/?q=es/download/subject_oriented&cl=en

GIL URDICIÁIN, Blanca. Manual de lenguajes documentales. Madrid (España) : Ediciones Trea , 2004.

GRACIA, Vilma. Términos de uso frecuente en las organizaciones. Coyuntura económica. En: <chrome://newtabhttp://www.google.com.co/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CCcQFjAB&url=http%3A%2F%2Fcoyunturaeconomica.com%2Fempresas%2Fterminologia-empresarial&ei=o4ypULiGBobK9gTWuYHYDQ&usg=AFQjCNGDtA9yJcf3JslMoy-ksDveSRBpuQ> Consultado en agosto 3 de 2015.

GUINCHAT, Claire y MENOU, Michel. Introducción general a las ciencias y técnicas de la información y documentación // Claire Guinchan y Michel Menou. 2 ed. España : : Cindoc, Unesco, , 1990.

ICONTEC NTC 1486. INSTITUTO COLOMBIANO DE NORMAS TECNICAS Y CERTIFICACION Normas Técnicas de Presentación de tesis, trabajos de grado y otros trabajos de investigación en. 6° Ed. 2008

ISO 25964-1:2011. Información y documentación - Tesauros y la interoperabilidad con otros vocabularios - Parte 1: Los tesauros para la recuperación de información. [Consulta: 07 de ene., de 2014]. Disponible en: http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=53657

ISO 5964:1985. Documentación - Directrices para el establecimiento y desarrollo de tesauros multilingües. [Consulta: 07 de ene., de 2014]. Disponible en: http://www.iso.org/iso/catalogue_detail.htm?csnumber=12159

LANCASTER, Frederick Wilfrid. Control del vocabulario en la recuperación de información. 2 ed. Valencia : Universidad de Valencia. 1995

LAMARCA LAPUENTE, María Jesús. Tesauros. Hipertexto: el nuevo concepto de documentos en la cultura de la imagen. En: <chrome://newtabhttp://www.google.com.co/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0CDMQFjAC&url=http%3A%2F%2Fwww.hipertexto.info%2Fdocumentos%2Ftesauros.htm&ei=i4-pUNu4JoTM9QSk64CADg&usg=AFQjCNErSwUVorm0qmJFz7Isu6hP7TfmxQ>. Consultado 9 de agosto de 2012.

MANOTAS ORTIZ, Nayibe. Utilización de la multimedia para el manejo de fuentes de información en una biblioteca Universitaria. Santafé de Bogotá : Universidad de la Salle , 1997

MICHAEL, Wilson. "Thesaurus relations as RDF schema - Example". In A Thesaurus Interchange format for migration to the Semantic Web. http://www.w3c.rl.ac.uk/pasttalks/slidemaker/XML_UK_SW_Thes/Overview.html

NAUMIS PEÑA, Catalina y TEXIA IGLESIAS MATURANA, María. Un lenguaje para la organización del conocimiento bibliotecológico en américa latina. En: dialnet.unirioja.es/descarga/articulo/1300407.pdf Consultado en septiembre 10 de 2012.

NC-ISO 2788: 2000. Directrices para el establecimiento y desarrollo de tesauros monolingües. [Consulta: 07 de ene., de 2014]. Disponible en: http://www.sld.cu/galerias/pdf/sitios/centromed/nc_iso_202788_directrices_para_el_establecimiento_y_desarrollo_de_tesauros_monolingues.pdf

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

PEÑA AYALA, Alejandro. Lenguaje Natural: Descripción de las etapas para su tratamiento. En línea: http://www.wolnm.org/apa/articulos/Lenguaje_Natural.pdf?target= [consulta: 07 de mar., de 2014]

RODRÍGUEZ LUNA, Cristina. Lenguajes documentales. España: Universidad de León. 2003. En: chrome://newtabhttp://www.google.com.co/url?sa=t&rct=j&q=yves+courier+define+enguajes+documentales&source=web&cd=1&ved=0CDEQFjAA&url=http%3A%2F%2Fwww.livu.com.ar%2Fliberatore%2Ffiles%2Farchivos%2Fdown%2FRodrigues%2520Luna_Lenguajes%2520documentales.pdf&ei=WOOuUMrsFoSg9QTVnYDQCA&usg=AFQjCNG9D-e31rtjTTKhb8Vb0ZwSXRhS4g Consultado en junio de 2012.

ROWLEY, Jennifer E. y HARTLEY Richard J.. Organización del Conocimiento: Introducción a la Gestión de Acceso a la Información. 4 ed. Aldershot, Inglaterra, Burlington, VT: Ashgate, 2008.

SLYPE, Georges van ; HÍPOLA, Pedro (traductor). Los lenguajes de indización : concepción, construcción y utilización de los sistemas documentales. Madrid : Fundación Germán Sánchez Ruipérez, 1991

SECTUR. Tesauro en el sector turístico e información relacionada con las actividades turísticas. Ministerio de Turismo de Argentina. [Consulta: 07 de ene., de 2014]. Disponible en: <http://www.cienaniosdeturismo.gov.ar/pagina.asp?IdSeccion=19>

TEMA TRES. Servidor de vocabularios controlados libre. [Consulta: 07 de ene., de 2014]. Disponible en: <http://www.r020.com.ar/tematres/>

THESAURUS de La Biblioteca de la Empresa Familiar. Carles Ferrer Salat. España. [Consulta: 07 de ene., de 2014]. Disponible en: http://www.biblio ferrersalat.com/index_saurus.php

UNE 50-106-90. Directrices para el establecimiento y desarrollo de tesauros monolingües. [Consulta: 07 de ene., de 2014]. Disponible en: http://elmemorioso.cl/une50106_90.pdf

UNE 50-121-91. Métodos para el análisis de documentos, determinación de su contenido y selección de los términos de indización. [Consulta: 07 de ene., de 2014]. Disponible en: http://elmemorioso.cl/une50121_91.pdf

UNESCO. UNESCO Thesaurus. <http://www.ulcc.ac.uk/unesco/>

VIZCAYA ALONSO, Dolores. Lenguajes Documentarios. Argentina : Nuevo Parhagigma Ediciones, , 1997.

ANEXO A

FORMATO DE EXTRACCIÓN Y ACTUALIZACIÓN DE TÉRMINOS PARA EL TESAURO ESPECIALIZADO EN EL AREA COMERCIAL

ANALISTA			
FECHA			
	DD	MM	AAAA
ACTUALIZACIÓN			ADICIÓN SUPRESIÓN CORRECCIÓN
TÉRMINO			PÁG. 2
TÉRMINO GENERAL (TG):			
ESPECÍFICO (EF):			
RELACIONADO (RE):			
LEADO POR (LP):			
LEE (LE):			
NOTA DE ALGUNDA:			
Marque con una "X" el documento técnico de cual se extrajo el TÉRMINO	Diagnóstico Técnico Concepto Técnico Estudio Técnico Otro documento	N° N° N°	
FUENTES DE VALIDACIÓN:			
OBSERVACIONES			
DEFINICIÓN			
FECHA DE APROBACIÓN			FECHA DE INCORPORACIÓN
DD/MM/AAAA			DD/MM/AAAA
RESPONSABLE			

Fuente: El propio autor.

ANEXO B

FORMATO PARA LA EVALUACIÓN DE SOFTWARE PARA LA AUTOMATIZACIÓN DE TESAuros

MODULO	ÍTEM	ESPECIFICACIÓN TÉCNICA	TEMATRES					THESAGE'S ENGLISH DICTIONARY AND TESAUROS					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 microtesauro en ambiente Web ó 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 Sistémicos 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 tesauro monolingües en idioma español?					X	X					X				
	17	La velocidad en el procesamiento de las búsquedas es?				X				X							X
Usabilidad	18	Permite establecer hipervínculo a los documentos afines al termino?					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			
OBSERVACIONES: Debido a las múltiples ventajas ofrecidas por TemaTres y producto de las ponderaciones, se decide considerar la incorporación de la estructura del microtesauro en esta aplicación, porque cumple con las necesidades del Centro de Documentación e Información de ésta entidad publica Distrital.			0	0	3	8	90	3	2	15	1	55	4	6	12	12	35
			96%					72%					66%				

Fuente: El propio autor.