

Génesis del tesoro documental como elemento de la representación temática y presencia actual

CATALINA NAUMIS PEÑA

En los veinticinco años que cumple el Centro Universitario de Investigaciones Bibliotecológicas han sucedido cambios importantes en la representación temática de los recursos bibliográficos y se avizoran muchos más en el panorama de la organización documental. El objetivo de este trabajo es hacer una revisión histórica, una prospección del panorama actual del mencionado fenómeno a través de proyectos recientes de investigación para desarrollar tesauros.

La humanidad ha experimentado en las últimas décadas grandes y aceleradas transformaciones económicas, políticas y culturales. Los cambios se produjeron en varios órdenes de las esferas sociales y personales tal como sucedió en la revolución industrial. La nueva revolución de la información o del conocimiento que se vive tiene consecuencias en las formas de trabajo, en la cultura y la universalización del acceso a la información.

La Bibliotecología tiene un papel trascendente en el proceso social de la revolución informática como mediadora entre las fuentes de conocimiento y la sociedad. En la medida que esa mediación se realice con responsabilidad y compromiso será piedra fundamental en el acceso a la información para mejorar la calidad de vida. La difusión del conocimiento depende de una organización documental breve, exacta, clara, consistente, uniforme y concisa, usando los adjetivos que proponía Escamilla para desarrollar los índices de las bibliografías.¹ La perspectiva bibliotecológica debe ayudar a enfrentar y afrontar la evolución y el progreso en la transmisión del conocimiento, para integrar

1 Gloria Escamilla (1998) *Manual de metodología y técnica bibliográficas*. P. 42, 45-46.

ciencia, tecnología, necesidades humanas y distintos puntos de vista culturales y lingüísticos.

La organización documental implica un conjunto de desafíos para adecuarse a los nuevos entornos tecnológicos, que suponen una respuesta oportuna a las necesidades de información de la sociedad. La información se transmite por diferentes medios de comunicación y llega a través de: radio, televisión, periódicos, libros, revistas, Internet, discos, videos, i-pod, películas. Tal cantidad de información no asegura que se obtenga conocimiento porque este es un proceso diferente. La transformación en conocimiento implica mantener una actitud analítica, sintética, crítica y reflexiva acerca de la información que se recibe y la contrastación con la experiencia anterior y las necesidades actuales. Por ello, la información que proporcione el bibliotecólogo deberá ser motivadora así como disparar los procesos de conocimiento.

La información no es conocimiento; existe un acto de conocimiento cuando la persona adquiere capacidad de discernimiento, es decir, selecciona y discrimina la información que realmente le interesa, realiza además un ejercicio constante de análisis y síntesis, de co-construcción y reconstrucción de los datos que se tienen, para contrastar con el propio marco referencial que el individuo posee, así la información se transforma en conocimiento útil, para el propio individuo y la sociedad en la cual está inserto.²

La puesta en marcha, desarrollo y mantenimiento de una serie de estrategias hará que la información se convierta en conocimiento. En primer lugar, el ser humano discrimina aquella información relevante y de su interés. Tras haber seleccionado la información, la analiza desde una postura reflexiva, intentando profundizar en cada uno de los elementos, deconstruyendo el mensaje, para coconstruirlo desde su propia realidad. Es decir en el proceso de deconstrucción se desmonta, comprende, entiende las variables, partes, objetivos, elementos, axiomas del mensaje. En el proceso de coconstrucción se realiza el procedimiento inverso. A partir de variables, axiomas, elementos, etcétera, se vuelve a componer el mensaje, desde la realidad perso-

2 Cisneros Rodríguez, I., García Doctor, C. y Lozano Jurado, M. (1999).

nal, social, histórica, cultural y vital, es decir, desde la perspectiva global del conocimiento y la persona.

En este proceso de conocimiento el bibliotecólogo cumple un papel trascendente porque será el mediador entre los diferentes soportes de información, sus contenidos y la sociedad, que es un rol de responsabilidad porque la mediación implica interpretación de los contenidos documentales, ya que no es una mediación automática como la que realiza la computadora, sino razonada para representar el pensamiento expresado en el documento. Para ello, es necesario entender el mensaje que se transmite, es decir convertir la información en conocimiento.

En el nuevo marco social de la información como valor económico, la biblioteca es además de un centro de acopio documental, un centro de distribución de documentos a distancia. Los sistemas de información especializados de texto completo brindan un servicio que por lo costoso aseguran cierto rigor en los documentos incluidos. En cambio, los documentos que se distribuyen por Internet no pasan por procesos de validación social como los artículos de revistas o libros indizados por sistemas de información o los sistemas de bibliotecas que operan en cierto modo como filtro de información, a través del apoyo de sus usuarios en las bibliotecas públicas o de los docentes o investigadores de las instituciones educativas, industriales o comerciales para asegurar que los documentos distribuidos mantengan cierto nivel de calidad. Si bien la WEB es dinámica en producción de información el resultado es incontrolado y esporádico. Los recursos informativos son volátiles y los servidores de las redes que los soportan cambian constantemente.

ANTECEDENTES DEL PROCESO DOCUMENTAL AUTOMATIZADO

En los años sesenta comenzaron a funcionar grandes centros de cómputo para procesar el trabajo bibliotecario y es en el año 1964 que una comisión del Ohio College Library Center (OCLC) comienza el primer proyecto bibliotecario de transferencia de información en línea, le sigue en 1966 la Library of Congress que inicia el proyecto

MARC (Machine Readable Catalog) para elaborar el catálogo en línea. Este proyecto contó con la colaboración de otras 16 bibliotecas de los Estados Unidos. Las bibliotecas inglesas también trabajaban en la producción de catálogos. Muchas bibliotecas mantenían catálogos impresos y usaban las computadoras para generarlos.

En los años setenta hubo un cambio cualitativo sustancial: se pasó del trabajo como tarea única a la operación de varios usuarios a la vez y además se inició el teleprocesamiento. Con la operación desde una terminal se podía conocer en cualquier instante la situación precisa de un libro o de un usuario, o de la colección como un todo. Sin embargo, al comienzo de la década de los setenta las computadoras seguían siendo aparatos extremadamente caros y complicados, que requerían forzosamente de todo un centro especializado en torno suyo. Era posible utilizarlas más ampliamente, pero su tiempo seguía siendo muy costoso, y para explotar las ventajas de operar en “tiempo real” era preciso además gastar en comunicación telefónica permanente. Sólo las grandes bibliotecas pudieron disponer de estos sistemas, porque para la mayoría, los costos eran prohibitivos.

A finales de los setenta comienzan a difundirse las minicomputadoras que continúan resultando demasiado costosas para las bibliotecas y los procesos tendían a hacerse fuera de los recintos bibliotecarios, privilegiando el trabajo de cómputo. El bibliotecario dependía de las empresas comerciales, las cuales los marginaban de efectuar desarrollos independientes y las limitaciones que implicaba esta situación hizo perder el control de algunos procesos, entre ellos la catalogación de grandes bases de datos que se realizaron con independencia de la normalización bibliotecaria. El resultado ya lo conocemos porque desembocó en una gran cantidad de información, pero con una recuperación difícil para la gran mayoría de los usuarios.

Con el tiempo los equipos se hicieron más pequeños, más baratos y comenzaron a aparecer sistemas como el *microisis* que permitían trabajo de programación del bibliotecario, o empresas que comenzaron a buscar la participación bibliotecaria y los servicios se hicieron bastante eficientes comparados con los ofrecidos por sistemas de recuperación que no seguían la normatividad bibliográfica. Los grandes sistemas bibliotecarios trabajan en redes compartidas que se van

adaptando a sus necesidades y con gran conectividad entre los diferentes departamentos y usuarios. Las bibliotecas pequeñas que no pertenecen a sistemas se debaten en la compra de programas de automatización a veces demasiado costosos y que ni siquiera responden a las necesidades de los usuarios o la habilidad del bibliotecario para encontrar soluciones informáticas adecuadas y fáciles de implementar. En paralelo a este proceso se ofrecen servicios de información especializados, educativos, industriales, comerciales, artísticos, cuyo costo asegura la distribución de información de calidad.

La incorporación de la tecnología en los procesos y servicios de las bibliotecas les ha permitido avanzar y enfrentar los desafíos para satisfacer las necesidades que presentan los usuarios tanto presenciales como remotos y promover “la biblioteca sin muros”. La nueva responsabilidad es diseñar instrumentos adecuados a las nuevas tecnologías que potencien la recuperación de información especializada.

En este sentido, el desarrollo de tesauros y vocabularios controlados para indizar y recuperar información conformados por estructuras de conocimiento y el lenguaje formal de diferentes áreas disciplinarias permiten trabajar en un nivel de síntesis de los contenidos documentales en los sistemas de información especializada. Desde hace varios siglos la Bibliografía y la Bibliotecología usaban mecanismos de ordenación temática de los documentos. Sin embargo, desde hace medio siglo aproximadamente, ha habido un cambio profundo y estos estudios se encuentran insertados en un proyecto social más amplio, que se relaciona con el análisis de comportamientos terminológicos en las diferentes lenguas y con soluciones informáticas.

La investigación lingüística surge como una necesidad de la industria del lenguaje para desarrollar sistemas automáticos de comprensión y producción de la lengua hablada y escrita, con la finalidad de ofrecer servicios de todo tipo regidos por la lengua. La posibilidad de desarrollar procesos automáticos mediados en la lengua de nuestra cultura es fundamental tanto para mantenerla, como para insertarse en el mercado global.

En este sentido, los estudios interdisciplinarios sobre mediación lingüística surgen en Estados Unidos en los años cincuenta cuando comienza la búsqueda para obtener computadoras capaces de traducir textos automáticamente de lenguas extranjeras al inglés, particularmente de revistas científicas rusas. Cuando aparece la inteligencia artificial en la década de los sesenta, la lingüística computacional se convirtió en una rama de la inteligencia artificial, tratando con el nivel de comprensión humana y la producción de los lenguajes naturales.

Para traducir un lenguaje en otro, se observó que uno tenía que entender la sintaxis de ambos lenguajes, al menos en el nivel morfológico (la sintaxis de las palabras) y las frases enteras. Para entender la sintaxis, se debe entender la semántica del vocabulario y la pragmática del lenguaje. Lo que empezó como un esfuerzo para traducir textos se convirtió en una disciplina encargada de entender cómo representar y procesar el lenguaje natural usando computadoras. Este tipo de investigación se conoce también como *lingüística* de corpus. A medida que se fue avanzando en las investigaciones en este campo, cada vez más se requirió de la participación de otras disciplinas que también estudian la mediación lingüística.

La Bibliotecología desde un principio intervino y fue parte de las investigaciones aportando elementos para lograr sistemas más eficientes de recuperación de información. La catalogación fue una de las primeras tareas asistidas por la computadora, para representar características de forma y contenido de los documentos, compartiendo créditos con el préstamo de libros en las grandes bibliotecas. La mediación lingüística se seguía realizando a través de las listas de encabezamientos de materia, pero utilizadas y aprovechadas de manera diferente a los sistemas manuales que habían prevalecido hasta la mitad del siglo XX, porque se combina con la indización automática de palabras en títulos y/o resúmenes, solución aportada por el campo de la computación. En estos primeros años también comenzaron los sistemas de recuperación de artículos de revistas en los cuales la mediación se realizaba a través de las palabras significativas de los títulos desde el reconocimiento automático, pero la Bibliotecología continuaba haciendo propuestas innovadoras como la construcción de te-

sauros para indizar documentos, incorporando de este modo, el aspecto humano y racional en la recuperación de información.

Estos antecedentes sobre los procesos documentales automatizados entremezclados con las soluciones bibliotecológicas nos introducen en la representación temática desde un punto de vista integrado, pero con un énfasis mayor en recuperación de información de calidad.

APARICIÓN EN ESCENA DEL TESAURO DOCUMENTAL

A lo largo de varios siglos han evolucionado los índices para representar temáticamente las obras, al interior de las propias obras o en las relaciones de obras sobre asuntos específicos. Desde el índice de materias agregado a su libro *De scriptis medicis libri duo* en 1662 por Juan Antonio van der Linden³ se han seguido refinando las técnicas de indización y sobre todo reflejando en ellos, los nuevos conocimientos que surgen.

Un índice no tiene valor separado de las obras que clasifica y en los registros de recursos bibliográficos, catálogos o bibliografías es un punto de acceso más a las obras. La importancia del conjunto de elementos de acceso al documento lo explica Napoleón I en una carta fechada el 19 de abril de 1807, cuando anunciaba la creación de la École des Chartes :

Si en una gran capital como París hubiese una escuela especial de historia donde se siguiera primero un curso de bibliografía, un joven, en vez de extraviarse durante meses en lecturas insuficientes o poco dignas de confianza, podría ir hacia las mejores obras y conseguiría más fácil y más rápidamente, mejor instrucción.⁴

En los siglos XIX y XX aparecen los sistemas de clasificación basados en los temas para ordenar los libros en las estanterías de las bibliotecas y los registros bibliográficos acompañados además por los encabezamientos de materia o epígrafes

3 L. M. Malclès (1960) *La bibliografía*. p. 28.

4 L. M. Malclès (1960) *Op. cit.* p.44.

En cuanto a la primera vez que el *tesauro* fue utilizado como un tipo de instrumento lingüístico en la indización de los documentos, Chaumier presenta en su libro la opinión de Vickery y Lancaster. Según Vickery, fue Helen Brownson de la American National Science Foundation, la primera que lo empleó en la: “Dorking Conference on Classification Research” (1957), presentando la “aplicación de un tesauro mecanizado basado en redes de significados relacionados”. En cambio, Lancaster le otorga el mérito a Hans Peter Luhn, quien lo empleó en un texto del mismo año 1957.⁵

Hans Peter Luhn fue científico de la computación para IBM, creador del algoritmo de Luhn, del concepto de Diseminación Selectiva de Información, de los resúmenes automáticos basados en la frecuencia estadística de términos y del sistema KWIC (Key Words In Context). Este emigrante alemán en Estados Unidos, hijo de impresores, trabajó en la *División de Investigación sobre Recuperación de Información en IBM*⁶. La trascendencia del trabajo de Luhn es difícil de opacar, pero en todo caso Helen Brownson adoptó y aplicó el concepto de tesauro con mucha celeridad.

La primera noticia sobre la palabra tesauro proviene del mundo helénico y se usaba para designar a las pequeñas capillas donde se guardaban los exvotos y las donaciones que frecuentemente eran muy ricas y valiosas, verdaderas joyas. Estas capillas estaban enclavadas en el recinto del santuario o témenos donde se encontraba el oráculo. El único tesoro restaurado que ha llegado a nuestros días pertenecía al templo de Apolo y la descripción bastante exacta de cómo fue el recinto sagrado se conoce gracias a las informaciones de Pausanias en el siglo II AC.

Posteriormente se usó para designar ciertos diccionarios, índices o repertorios de términos,⁷ con la misma connotación de lugar donde se guarda lo más importante, que en este caso es menos corpóreo

5 Jacques Chaumier (1986) *Análisis y lenguajes documentales: el tratamiento lingüístico de la información documental*. p. 11.

6 Wikipedia: the free encyclopedia (Consultada el 12 de septiembre de 2006) http://en.wikipedia.org/wiki/Hans_Peter_Luhn

7 J. Martínez de Souza (1993) *Diccionario de bibliología y ciencias afines*. – p. 776.

porque se trata del saber humano. A lo largo de la historia, el término tesaurus está asociado a obras de estructuras muy variadas tanto diccionarios como repertorios terminológicos o vocabularios. De hecho, la voz *thesaurus* en latín fue, durante años, el término más utilizado para designar tanto un diccionario muy exhaustivo (de lenguas muertas) como un vocabulario que contuviera todas sus unidades léxicas:

En la actualidad, la voz de *tesaurus* en el ámbito bibliotecológico se usa en el sentido de redes de significados relacionados, tal como lo planteó Helen Brownson en 1957, aunque sigue apareciendo en algunos diccionarios y enciclopedias usado en el valor de diccionario, catálogo o antología.⁸ El *Oxford English Dictionary* define el *tesaurus* como “una lista clasificada de términos, especialmente términos llave, en un campo en particular, para usar en la indización y recuperación de información”⁹

En 1969 Aitchison publica el *Tesaurus facetado* que muestra las ventajas de indizar y clasificar con una correspondencia entre facetas y descriptores (Gilchrist, 2003: 9) En un principio los sistemas de indización y recuperación que utilizaban el tesaurus trabajaban con tarjetas perforadas en lotes, pero luego se aplicó a sistemas de recuperación en línea, con éxito. La permanencia de la herramienta determinó la propuesta de normas para construirlos (*International Organization for Standardization 2788*, 1986). Incluso en el año 2000 Aitchison y Gilchrist publican un manual para elaborar tesaurus. La National Information Standards Organization en el año 2005 publica las “Guidelines for the construction, format and management of monolingual controlled vocabularies”. Estas últimas normas agregan las taxonomías como herramientas a ser desarrolladas con su apoyo y sobre todo involucra la organización de información electrónica.

En el campo de la literatura el tesaurus es una lista de palabras con significados similares sinónimos, habitualmente acompañada por otra lista de antónimos. Un ejemplo sería un tesaurus dedicado a un

8 *El pequeño Larousse ilustrado 1996 en color*. – México: Ediciones Larousse. – 1792 p.

9 *The Oxford English Dictionary*. Vol 17, p. 924.

campo especializado, que contiene la jerga que se emplea en dicho campo del conocimiento. En el mundo de habla inglesa, es clásico el Tesauro de Roget´s cuya función es, según su autor, además de ayudar al escritor a encontrar la palabra que exprese mejor su pensamiento, también estimular su intelecto y sugerirle palabras o ideas relacionadas.¹⁰ En cierto modo, en el tesauro que usan los documentalistas y en el Tesauro de Roget existen ciertos objetivos similares, pero difieren en el uso y en la presentación.

El término tesauro es muy usado para denominar diccionarios del tipo del Roget´s, por esta razón se prefiere el uso del término tesauro documental hacia el medio social, como se indica en el nombre del trabajo, ya que el uso de un término compuesto evita la polisemia que lo acompaña, aunque al interior de la Bibliotecología la referencia es clara. Si bien, la designación del nombre tesauro puede resultar motivo de discusión, en la literatura bibliotecológica se destaca como tema prioritario la propuesta de grandes proyectos de construcción de tesauros, en los cuáles se observa un trabajo de investigación minucioso y costoso

LA CONSTRUCCIÓN DEL TESAURO DOCUMENTAL DESDE UN PARADIGMA CIENTÍFICO

El tesauro mantiene una estructura agrupada en campos semánticos y su clasificación temática se representa a través de términos normalizados y validados como descriptores, dotándolos con características propias y adecuadas a las necesidades del sistema de información del que forman parte.

La construcción de un tesauro debe ser resuelta a la luz del método científico puesto que deben ser probadas hipótesis, analizadas variables, aplicados métodos cualitativos y cuantitativos que deriven en un resultado que metodológicamente pueda ser repetible en otros lenguajes o entornos científicos.

10 Roget´s Thesaurus. p. xviii.

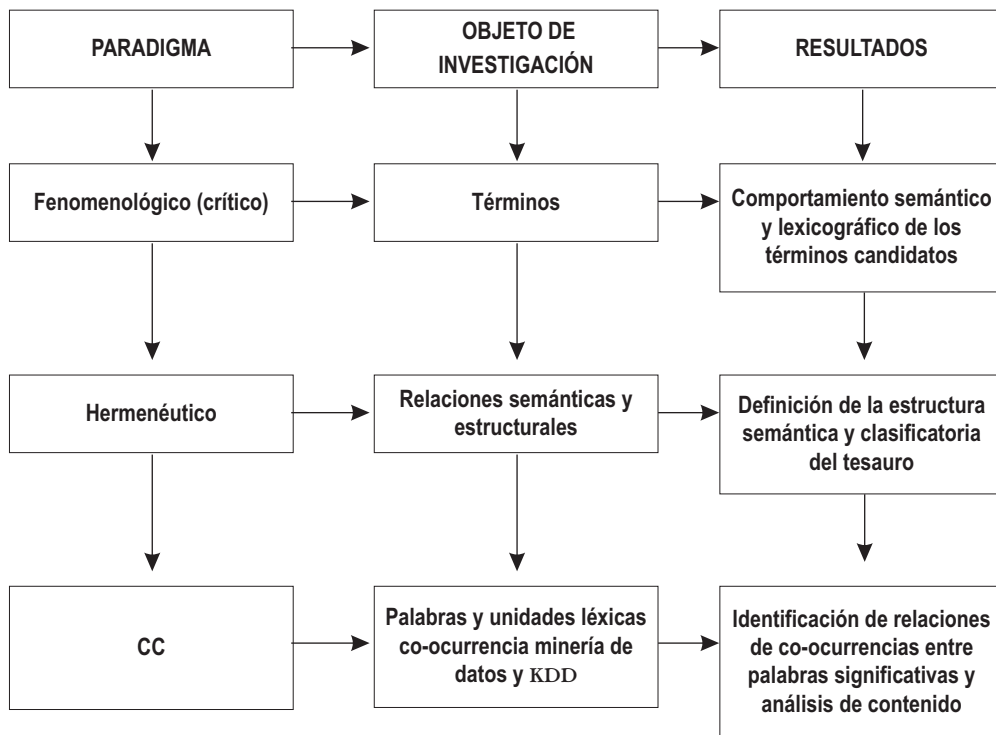
Lo anterior se refuerza si analizamos como en la construcción de un tesaurus pueden estar presentes diversos paradigmas tales como el fenomenológico o crítico, (cuando cada término es sometido a un juicio crítico no sólo desde el punto de vista semántico sino también lexicográfico); el cuantitativo (cuando la asignación de los términos que intervienen en el tesaurus son previamente evaluados o proporcionados a partir de un estudio métrico de co-ocurrencia de palabras o de términos en una base de datos textual, minería de datos o técnica de Knowledge Discovery Data Base, por ejemplo) o el hermenéutico (cuando la estructura semántica obtenida para definir el tesaurus es el resultado de un minucioso procedimiento de clasificación y comparación). De acuerdo con el uso de alguno de estos paradigmas en el proceso o desarrollo de la investigación el producto final obtenido podrá hacer mayor énfasis en la calidad terminológica, en su estructura semántica o en la certeza o fiabilidad de las relaciones lexicográficas que se propongan. Las relaciones paradigmáticas anteriores pueden ser mostradas en el modelo de investigación que se presenta en la siguiente página.

La investigación para elaborar un tesaurus supone diferentes combinaciones de variables que se deben contemplar entre las cuáles se destacan:

- ❖ Comprensión y estudio de la temática.
- ❖ Estudio del corpus terminológico.
- ❖ Definición de los mapas conceptuales o árboles de dominio de un ámbito de conocimiento.
- ❖ Identificación de las características del medio de transmisión de información.
- ❖ Búsqueda de equivalencias entre las expresiones de un idioma a otro.
- ❖ Búsqueda de equivalencias entre términos del mismo idioma.
- ❖ Reconocimiento de las relaciones semánticas entre términos.
- ❖ Evaluación de las necesidades de los usuarios.
- ❖ Identificación de las expresiones que espera el tipo de usuarios que consultará el sistema de investigación.

Figura 1

Modelo de investigación orientado a la elaboración de tesauros documentales



Cada una de las variables mencionadas responden a estudios que exigen rigurosidad para analizarlos e interrelacionarlos entre sí, como un conjunto que conformará el tesoro.

En la construcción de herramientas lingüísticas para buscar y recuperar información temática se han ido presentando situaciones de prueba y error que mejoran los sistemas, pero las soluciones más contundentes provienen de la investigación donde se controlan las variables, se observan y se analizan los resultados contrastándolos con la práctica. Los controles de autoridad de temas en los servicios de información de ámbitos especializados pueden estar implementa-

dos de diferentes maneras, pero se destacan entre algunas de las soluciones, los lenguajes controlados definidos en la experiencia o aquellos producidos con la experiencia acumulada en los servicios y sometidos a un proceso de investigación que le otorga calidad a la validación de los términos utilizados y lo que es más importante, estudia y define la situación comunicativa que ocurre. El conocimiento del proceso comunicativo culmina con un diseño para indizar y recuperar los contenidos documentales.

El conjunto de necesidades sociales y de prácticas a dilucidar en torno al tema ha preocupado a los bibliotecólogos para proponer soluciones. En este sentido, con el tesoro documental mediante términos y descriptores, se pretende describir documentos por el contenido, destinados a usuarios especializados en una rama del conocimiento. Se discutirán a continuación los factores que inciden en la confusión entre investigación y práctica profesional.

INVESTIGACIÓN Y PRÁCTICA PROFESIONAL

Para estandarizar la representación y comunicación de la información los organismos internacionales reguladores de la actividad bibliotecológica han elaborado normas con la finalidad de construir tesauros, que deben ser cumplidas para asegurar la armonía entre diferentes sistemas de información que permitan el intercambio y la opción de utilizar programas que los respalden. Sin embargo en la construcción de un tesoro, el problema no son las normas que se aplican, sino el conjunto de elementos listados en el apartado anterior, que varían según la situación comunicativa a resolver. ¿Se puede decir que el proceso de trabajo es distinto pero con idénticos fundamentos? ¿En cada práctica existe una concepción distinta del tesoro?

Al desarrollar un tesoro se trabaja en un campo de conocimiento, cuyo objeto circunstancial son las unidades terminológicas, el marco es la representación y comunicación especializada y el vehículo son los documentos. La teoría supone que el objetivo del tesoro es el análisis y la explicación del fenómeno de la representatividad y comunicación en función de la finalidad y objetivos de la situación comunicativa

particular. La práctica resuelve cuestiones de representación y comunicación en términos generales, mediante el reconocimiento de las unidades terminológicas y los documentos que las contienen.

Dos situaciones que crean confusión con respecto a la investigación en torno al tesoro son las normas que regulan su construcción, un objeto, un marco y un vehículo que siempre parecerían ser los mismos. Las normas regulan muchas cosas en la vida social y son útiles para evitar conflictos y tratar de mantener un equilibrio de circunstancias que pueden tener similitudes. Sin embargo, la misma aplicación de la norma implica el estudio de las circunstancias específicas involucradas en la situación a resolver, no sólo para definir las normas y la forma de aplicación, sino para analizar y explicar el fenómeno o situación ocurrida. La existencia de un objeto, el marco y el vehículo que parecerían ser los elementos definidos de la investigación, suponen que no hay nada que investigar y se debe únicamente resolver algunas cuestiones implicadas en el proceso de representación y comunicación especializada.

La existencia de un objeto preciso y unificado que son los términos y de un vehículo que es el documental, no solucionan el problema de la representación y comunicación temática en un ámbito especializado, porque resolverlo supone la obtención de datos reales y representativos del universo a describir, y la necesidad de reflexión y definición sobre los puntos de observación. Incluso se puede utilizar una metodología similar para definir la estructura de conocimiento a representar en el tesoro, los términos y las relaciones entre ellos, pero siempre está implícita la observación, la descripción y la explicación de los términos, la estructura de conocimiento y las relaciones terminológicas representativas de una especialidad, más las necesidades del usuario. El documento como vehículo también presenta una gama bastante amplia de posibilidades porque no sólo se relaciona con la variación en el tipo de documento, sino que la situación comunicativa particular modifica el estilo y la estructura de los textos y por lo tanto, se deben analizar las repercusiones en la transferencia temática.

Los términos tienen un componente cognitivo, un componente lingüístico y otro sociocomunicativo que deben ser examinados para

definir la tipología de los nodos y las relaciones entre ellos. Los descriptores del tesaurus ponen el énfasis en los componentes cognitivos y sociocomunicativos del término que le sirvió de base para representar y transferir información. Las variaciones discursivas, tanto denominativas como conceptuales, tienen que ver con las condiciones de producción, transmisión y recepción. Una variación denominativa casi siempre tiene consecuencias cognitivas porque reflejan puntos de vista, escuelas de pensamiento, posiciones ideológicas, intereses económicos, y algunos elementos adicionales que se nos pueden estar escapando. Otros problemas terminológicos importantes que deben ser observados y analizados tienen que ver con la denominación de conceptos especializados que comparten la representación y comunicación de conceptos en otras especialidades o dentro de la misma especialidad. El descriptor busca la univocidad del término escogido para transferir el contenido documental al medio social, aunque siempre es necesario valorar las consecuencias cognitivas que trae aparejada.

Todo campo de conocimiento permite tratar a su objeto desde tres perspectivas: teórica, aplicada y como aplicación en las necesidades socio-profesionales. Todo campo de conocimiento posee un objeto central que trata de describir, explicar y contrastar desde un punto de vista cualitativo y cuantitativo. El objeto de estudio en la Bibliotecología, y más específicamente en la construcción de cada uno de los tesaurus que se aplicarán en la práctica profesional, es representar y comunicar los conceptos esenciales de los textos especializados a través de descriptores para facilitar su recuperación por el contenido, poniendo énfasis en la intermediación del documento con el usuario.

Cada tesaurus es un problema de representación y transferencia de información particular cuando se trata de vincular un sistema de información con usuarios especializados en un ámbito de conocimiento. La observación de la práctica profesional, el estudio de los antecedentes, una revisión bibliográfica a fondo, el análisis para transformar el planteamiento de forma más precisa y estructurada de los factores que inciden en la representación y la comunicación, la selección de la perspectiva principal desde donde se abordará la estructura del tesaurus, los estudios estadísticos de los términos, más la

aplicación de métodos de investigación cualitativos, darán como resultado un tesoro objetivo, útil y que refleje una estructura que permita representar y comunicar los documentos de una especialidad.

Cualquier campo de conocimiento en su vertiente aplicada tiene objetivos y una necesidad a cubrir, por lo que concibe su objeto para determinar los criterios de pertinencia. En este sentido desde la práctica profesional se pueden recopilar los términos usados en los documentos especializados y se pueden agregar las equivalencias con sinónimos que vayan apareciendo en la comunicación, para representar los documentos que se incorporan al sistema de información y se pueden aplicar algunas de las normas que regulan la construcción de tesoros. Sin embargo, las responsabilidades de la práctica profesional difícilmente permiten desarrollar una investigación objetiva para fundamentar la estructura conceptual de la especialidad con la categorización de sus conceptos y estudiar la situación comunicativa con la finalidad de proponer un tesoro eficiente.

Otro de los aspectos cuestionados en la construcción de un tesoro es la base práctica desde la cual se instrumentan los descriptores y los mapas conceptuales que representan la estructura de conocimiento subyacente en los sistemas de información. Charles A. Cutter definió el *principio del uso* como uno de los fundamentos del vocabulario controlado.¹¹ Durante muchos años fue interpretado como el término más correcto en el lenguaje escrito, sin embargo en la comunicación actual se prefiere el término más habitual entre los usuarios finales. Los documentos interpretados pueden ser videos que reflejen los usos populares de los términos, por lo cual los términos preferidos serán términos populares.

Es necesario aclarar que, la representación por términos populares no significa que un tesoro no esté basado en un trabajo de investigación, tomando en cuenta que además deberán incluirse las otras formas de presentación de la equivalencia terminológica. Los mapas conceptuales y su definición son una preocupación reflejada en la literatura de la especialidad. Sin embargo, son pocos los trabajos de

11 Foskett (1996).

investigación que estudian en profundidad la estructura de un ámbito del conocimiento, quizás por la dificultad de expresar situaciones que ocurren en el nivel cognitivo de la mente humana. La conformación de los mapas conceptuales conlleva el conocimiento profundo de las diferentes posibilidades para estructurar la más convincente, con la finalidad de satisfacer las necesidades organizativas de un ámbito especializado. También es cierto que cada vez más tesauros se construyen sobre la base de una investigación que plasme la organización que requiere un ámbito del conocimiento y la situación comunicativa en particular.

La literatura bibliotecológica divide los tesauros documentales en: monolingües y multilingües por la lengua en la que están contruidos; monodisciplinarios, multidisciplinarios e interdisciplinarios por la diversidad e interrelación de los temas que incluyen. La tipología retoma un aspecto o un comportamiento que suministra un nivel de abstracción para definir las clases del término a clasificar y nos permite establecer diferentes estudios en ese nivel de abstracción. La aplicación del paradigma hermenéutico, que conduce a la división semántica, no supone la restricción en una interacción con otros niveles y otras relaciones en el mismo nivel semántico.

El tesoro documental es motivo de estudio en su parte idiomática porque el *corpus* lingüístico representado es un modelo a escala de la variedad de la lengua en un determinado ámbito de conocimiento. En este sentido existe una definición de *corpus* que precisa el concepto de representatividad, que es el “equilibrio conceptual” o “balance conceptual” que supone recoger todos los subcampos en los que se divide cada parte del conocimiento, así como todos los ámbitos especializados que guardan relación con ésta (si es que se trata de un ámbito multidisciplinar). De acuerdo con este concepto, parece más fácil que el *corpus* sea representativo si incluye todas las variedades textuales típicas del área como objeto de estudio y aborda todos los subcampos en

los que ésta se divide y aquellos con los que guarda relación. El concepto de *equilibrio conceptual* fue introducido por Bowker.¹²

Cada tesoro que se construye responde a un sistema de información y aunque el área disciplinaria y el idioma sean los mismos, las concepciones son diferentes y las expresiones varían de una región a otra. En un estudio comparativo de tesauros de Bibliotecología en España y México se concluye que:

[...] la diferencia fundamental que se aprecia entre los dos tesauros es una estructura disciplinar identificando los fenómenos de constitución de las categorías semánticas, en ambos casos heredada de escuelas y teorías diferentes: teoría anglosajona Library and Information Science y Teoría del documento o de la Documentación de la escuela belga encabezada por Paul Otlet.¹³

La investigación para elaborar tesauros multilingües es otro reto, ya que cada lengua responde a un sistema de representación. Los lingüistas desarrollan diccionarios especializados con equivalencias a otras lenguas. Sin embargo, cuando se indiza con tesauros elaborados en una lengua diferente a la del sistema de información, los diccionarios son insuficientes porque tardan en reflejar los usos lingüísticos y la investigación de las equivalencias insume mucho tiempo. Los indizadores no deben perder tiempo en realizar la traducción de los términos y debe existir un equipo de investigación que defina la conveniencia de preferir una equivalencia sobre otra con base en un paradigma fenomenológico.

Las grandes bases de datos tienen equipos de apoyo para traducir los documentos y para definir los términos de indización más representativos de los documentos en diferentes idiomas.¹⁴ Se pueden mencionar además las investigaciones lingüísticas y la producción de vocabularios especializados que crecen día a día en el mundo entero, sobre los cuales debemos estar alertas, porque su conocimiento es

12 L. Bowker (1996) "Towards a Corpus-Based Approach to Terminography" p. 45-46.

13 C. Naumis (2007) "Estudio comparativo de tesauros bibliotecológicos en lengua española"

14 G. Rosemblat y L. Graham (2006) "Cross-language search in a monolingual health information sistema: flexible designs and lexical processes". p. 173.

básico para desarrollar nuestra propia investigación. El conocimiento de estas obras demuestra que además de la necesidad del estudio de las equivalencias idiomáticas, los enfoques sobre determinado tema en general difieren y que debemos realizar nuestras propias investigaciones contextualizadas en un determinado sistema de información para representar información y recuperar contenidos.

PROYECTOS DE INVESTIGACIÓN RELACIONADOS CON TESAUROS

Se realizó una búsqueda de proyectos de investigación sobre tesauros y se consideraron como fuente válida todas aquellas investigaciones publicadas en revistas de alto impacto o proyectos publicados en la *www* que especifiquen procedencia, autoría y representatividad de un organismo oficial de conocido prestigio internacional como universidades o entidades gubernamentales acreditadas.

A partir de la recopilación se observó cómo se trata de asegurar la interoperabilidad entre vocabularios controlados (incluyendo vocabularios controlados y no-controlados) y sistemas de organización del conocimiento; se observaron además esfuerzos para establecer interoperabilidad entre vocabularios en el mismo lenguaje o en diferentes lenguajes, entre diferentes sistemas de clasificación y entre vocabularios controlados y esquemas de clasificación.

Estos trabajos se han conducido hacia el mapeo e integración de sistemas de organización del conocimiento existentes o a la creación de nuevos, para compartir información en un medio de redes. Los proyectos varían en dos sentidos el mapeo y los métodos usados en el enfoque de su desarrollo, para ordenar la información en la red, desde los tradicionales sistemas de clasificación y tesauros, hasta las más novedosas taxonomías, ontologías y redes semánticas.¹⁵ En relación a las ontologías no se incluyen los proyectos que hacen hincapié

15 A. Martínez, C. Ristuccia, *et al.* (2004) "Las categorías o facetas fundamentales: una metodología para el diseño de taxonomías corporativas de sitios web argentinos".

en el software que las sustenta, porque en este trabajo interesa el ámbito lingüístico.

Se destaca además un interés creciente, tanto en organismos gubernamentales, internacionales y académicos, en la elaboración de herramientas lingüísticas y en la búsqueda de la interoperabilidad entre vocabularios con diversas clasificaciones y diferentes lenguajes. En todos los casos se aprecia un trabajo de investigación para definir la situación comunicativa, que en cada uno tiene una naturaleza diferente.¹⁶

A continuación se comentan, en un apartado previo al análisis, los últimos trabajos realizados relacionados con el tema.

INVESTIGACIÓN EN EL CUIB

Cuando hace veinticinco años se crea el CUIB, el grupo de iniciadores que provienen de la Dirección General de Bibliotecas de la UNAM habían vivido la experiencia de poner en marcha el sistema bibliográfico automatizado LIBRUNAM, con un equipo Burroughs 6700, la catalogación en línea en el Departamento de Procesos Técnicos y consulta de los catálogos de microfichas en los servicios bibliotecarios de la Universidad. Se preparó también la migración en 1985 al equipo Briton Lee- Alpha Micro. La representación temática de los contenidos documentales se realizaba y se realiza a partir de listas de encabezamientos de materia porque se trata de un catálogo centralizado universitario. Sin embargo, este tipo de herramienta también es utilizado en las innumerables bibliotecas especializadas y en las bibliotecas digitales actuales que exigen una representación más puntual de los contenidos. La investigación en el CUIB en un comienzo se dirigía a la representación temática con encabezamientos de materia.

Sólo en algunas bibliotecas especializadas de la Universidad Nacional Autónoma de México (UNAM) y para operar bases de datos de fondos especializados se han utilizado vocabularios controlados y

16 L. M. Chan y M. L. Zeng (2002) "Ensuring interoperability among Subject Vocabularies and Knowledge Organization Schemes: a methodological analysis".

tesauros para representar los contenidos temáticos de los documentos. En el CUIB se ha generado primero un vocabulario controlado para organizar los temas de la base de datos INFOBILA que contiene los documentos bibliotecológicos generados en América Latina y con posterioridad un tesoro con la misma finalidad, pero ampliando la cobertura de los términos a los usos de cuatro países latinoamericanos que fueron los participantes en su desarrollo: Chile, Colombia, México y Venezuela. Este mismo Centro ha desarrollado además un vocabulario controlado en Derechos Humanos y un tesoro para organizar los contenidos educativos que se transmiten por la Red EDUSAT, sistema que contempla la educación a distancia de educación básica, media y media superior, así como la difusión de programas culturales. En la UNAM no existen proyectos de investigación para desarrollar tesauros aunque se hacen investigaciones relacionadas con el tema y vocabularios controlados, fruto de la experiencia en la elaboración de registros bibliográficos.

En el CUIB se acaba de concluir una tesis de maestría sobre comportamientos informétricos en la representación de contenidos en ciencias bibliotecológicas y de la información relativa a América Latina. Se utiliza una técnica basada en la co-aparición de palabras y de términos extraídos de los contenidos de los documentos para identificar un conjunto de relaciones presentes y poco visibles en el sistema de Comunicación Científica. Entre los resultados obtenidos se derivó un mapa conceptual para representar la estructura semántica de la especialidad. Se observó cierta discrepancia entre cada una de las unidades léxicas estudiadas.¹⁷

En otra tesis de maestría se analiza el lenguaje científico sobre el tema del agua y la necesidad de desarrollar un tesoro en español para organizar una información prioritaria en la investigación actual.¹⁸ Un tercer trabajo, también de nivel maestría, estudia la terminología sobre género. En este último campo existen varios tesauros

17 M. P. López Rico (2006) *Comportamiento informétrico en la representación de contenidos en ciencias bibliotecológica y de la información relativo a América Latina*.

18 V. E. Vargas Suárez (2007) *Análisis del lenguaje científico sobre agua*.

en español, pero se detectan aún insuficiencias terminológicas y un mapeo insuficiente de los conocimientos que se manejan en el tema. En el CUIB, una de las investigaciones en curso analiza la terminología de la Bibliotecología y los nuevos términos que han ido apareciendo desde la publicación del tesaurus sobre el tema en el año 1999.

PROYECTOS ACADÉMICOS INTERNACIONALES

La Universidad de Bar-Ilan en Israel y la Fundación de Ciencia Israelita apoyan un proyecto de investigación para desarrollar un mapa de conocimiento de la Ciencia de la Información apoyado en un estudio teórico sólido. El proyecto está insertado en otro mayor de clasificación de las ciencias en general que en los dos últimos años ha estado dedicado, la mayor parte del tiempo, al proyecto específico que se comenta porque es el más documentado. El trabajo científico está basado en el Modelo Critical Delphi que es una metodología de investigación cualitativa orientada a facilitar las discusiones moderadas y críticas entre expertos (el panel). El panel internacional e intercultural está compuesto por 57 participantes provenientes de 16 países. Es único y excepcional, ya que abarca académicos de primer nivel quienes representan a casi todas las principales subáreas y aspectos importantes del área. Las discusiones indirectas fueron anónimas y llevadas a cabo en etapas sucesivas a través de cuestionarios estructurados.

El primer cuestionario contenía 24 preguntas abiertas y detalladas que abarcan 16 páginas. El segundo cuestionario contenía 18 preguntas en 16 páginas. El tercer cuestionario contenía 13 preguntas en 28 páginas. Los porcentajes de devolución fueron relativamente altos: 57 académicos (100%) devolvieron la primera etapa, 39 (68.4%) devolvieron la segunda etapa y 39 (68.4%) devolvieron la tercera etapa. Cuarenta y tres panelistas (75.4 %) participaron en dos etapas (T1 y (T 2 o T 3)) y 35 panelistas (61.4%) participaron en las tres etapas. Se devolvieron las respuestas a cada miembro del panel con reflexiones críticas relevantes. Cuarenta y siete participantes

(82,4%) respondieron y aprobaron sus respuestas. Veintitrés de ellos, los cuales representan un 48.9% (23 de 47), y 40.3% del panel completo (23 de 57) revisaron sus respuestas originales. Por lo tanto, se puede decir que el proceso crítico se compuso de cuatro etapas.

El siguiente paso fue agrupar las concepciones sobre Ciencia de la Información que difieren principalmente en tres temas clave, según lo establecido en las discusiones del panel, ellos son: fenómenos, dominio y alcance: ¿Cuáles son los fenómenos explorados? ¿Cuál es el dominio del área? ¿Cuál es el alcance de la exploración?¹⁹

La Universidad Estatal de Portland conjuntamente a la Fundación Nacional para la Ciencia de los Estados Unidos de Norteamérica desarrolló entre los años 2005 y 2006 el proyecto “SGER” para la indización acelerada en un dominio específico de conocimiento organizado en una biblioteca digital. Este tuvo un carácter exploratorio, cuyo foco de atención era el proveer de manera rápida el acceso a documentos de un área específica, en este caso relacionada con el manejo de recursos forestales. Los académicos adscritos al proyecto escribieron, indizaron y buscaron documentos utilizando terminología especializada (incluyendo en ocasiones palabras de uso normal en el inglés).

La valoración del sistema, al ser manipulado tanto por usuarios expertos como inexpertos, permitió explorar las estrategias de búsqueda, y seguir de cerca el uso del vocabulario controlado que se estaba proponiendo, así como el esquema de clasificación que complementaba los índices de los documentos. Dada la dificultad en la recuperación de la información en general, el proyecto fue considerado de alto riesgo, pero podría ser exitoso debido a que el trabajo se efectuaba en un dominio de conocimiento específico con una terminología bien controlada. El mérito intelectual de esta propuesta estribó en el uso de un nuevo modelo de tesoro “Metadata++”, con una representación de entradas basadas en términos consensuados con la indización. El principal impacto de este proyecto fue la ubicación del problema y la ayuda para la toma de decisiones con base en la infor-

19 Chaim Zins <http://www.success.co.il/knowledge/about.html>

mación recuperada, lo que coadyuvó a extender la colaboración entre el Departamento de Servicios Forestales, la organización no gubernamental “Nature Conservancy” y la biblioteca de Ciencias de la Información de la Denmark’s Royal School.²⁰

El Centro de Investigación para Bibliotecas Digitales de la Universidad de Stratchlyde, en Inglaterra, ha desarrollado una serie de proyectos denominados “High-Level Thesaurus”. Este surge como respuesta a la necesidad de conocer el uso de los distintos esquemas temáticos y su asociación con las prácticas de recuperación de información entre archivos, museos, bibliotecas, centros comunitarios de servicios de información electrónica, para satisfacer las necesidades de información de sus usuarios. Existe un gran consenso entre los profesionales de la información que participan en este proyecto sobre la necesidad del uso de rutas interactivas de las terminologías utilizadas entre diferentes comunidades, que permitan su interoperabilidad en los procesos de recuperación de información. El proyecto se encuentra actualmente en su tercera fase, que abarca de julio del 2006 a julio del 2007, cuyos objetivos son la investigación orientada al desarrollo de soluciones piloto a los problemas detectados en la búsqueda multi-temática a través de varios esquemas de información en diversos ambientes, así como la provisión de diversas ayudas de orden terminológico. El proyecto se apoya en una gran variedad de terminologías como son: *Art and Architecture Thesaurus (AAT)*, *The J. Paul Getty Trust*, *Dewey Decimal Classification (DDC)*, *OCLC Global Change Master Directory (GCMD)* (Science Keywords), *NASA HASSET Thesaurus*, *UK Data Archive at the University of Essex Integrated Public Sector Vocabulary (IPSV)*, *e-Government Unit (UK)*, *Joint Academic Coding System (JACS)*, *Universities and Colleges Admission Service (UK)*, *JITA Classification Schema*, *E-Prints in Library and Information Science (E-LIS)*, *Library of Congress Subject Headings (LCSH)*, *Library of Congress (USA)*, *Medical Subject Headings (MeSH)*, *National Library of Medicine (USA)*, *National Monuments Record Thesaurus (NMR)*, *English Heritage*,

20 Project Profile (2006) “SGER Accelerated Indexing in a Domain-Specific Digital Library”.

*UNESCO Thesaurus, UNESCO and the University of London Computer Centre.*²¹

PROYECTOS GUBERNAMENTALES

La Comisión Económica Europea desarrolló el proyecto “OntoGov”, con duración de 30 meses a partir del primer día del mes de enero de 2004. La idea principal era desarrollar el acercamiento (y un sistema) que soporte la composición, reconfiguración y evolución de los servicios e-gobierno de los países integrantes de la Unión Europea. Se conformó un consorcio de patrocinadores y desarrolladores tecnológicos (la consultora griega “Planet”, el Centro de Investigación afiliado a la Universidad de Karlsruhe, la empresa griega desarrolladora de tecnologías de la información “Archetypon”, la compañía española de servicios de alta tecnología “Indra”) y un grupo piloto de usuarios (Cancillería Federal de Suiza, la Municipalidad de Barcelona, España y la Municipalidad de Amaroussion, Grecia).

Las metas del proyecto eran: 1) ser el puente entre la brecha que existe entre la toma de decisiones y la realización técnica de los servicios de e-gobierno, 2) considerar el ciclo de vida de los servicios de e-gobierno, 3) hacer explícito el manejo de conocimiento. Los beneficios de los resultados esperados redundarían en establecer a “OntoGov” como la plataforma en la cual se basaran los diseños y rediseños de los servicios de e-gobierno, de tal forma que los cambios que se efectuaran no rompieran la consistencia de los servicios, así como la automatización del manejo de la administración pública que impactará en la formulación de una nueva legislación que contemple un espectro más amplio de los servicios de e-gobierno, que modifique solamente las partes del servicio que lo requieran y asegure la consistencia jurídica de los nuevos servicios.²²

21 High-Level Thesaurus (2006). University of Strathclyde, UK.

22 Ontology-enable e-Gov Services Configuration, (2006) Project IST-507232. European Comisión, Information Society Technologies.

Los gobiernos de once países europeos patrocinan el desarrollo de un tesoro multilingüe para organizar la información del patrimonio arqueológico y arquitectónico de los países participantes en el proyecto. El *Herein Thesaurus* se construyó sobre la base de los reportes de políticas del patrimonio cultural provenientes de España, Francia, Hungría, Irlanda, Noruega y Reino Unido. Los tres países que comenzaron a desarrollar las listas de términos en sus propias lenguas fueron España, Francia y Reino Unido. Posteriormente se desarrollaron las equivalencias y se establecieron las clases principales. Luego se fueron haciendo equivalencias al búlgaro, polaco, rumano, griego, lituano y esloveno. El esfuerzo en este trabajo estuvo centrado en la construcción de un instrumento común donde todos los lenguajes tienen un valor igual y se respetaron los términos específicos de cada cultura. Por ejemplo, lo significativo que es para la sociedad argentina el fútbol, llevó al Ministerio de Educación, Ciencia y Tecnología de la Nación en la República Argentina a subsidiar un proyecto sobre “Fútbol, tesauros y taxonomías Web: desafíos del control del vocabulario”. Se analizó la terminología empleada en las taxonomías de 31 sitios Web de clubes de fútbol argentinos contrastándolos con las normas de construcción de tesauros y las directrices para controlar autoridades de nombres y materias. No se hace una propuesta concreta de tesoro, pero se concluye que se puede construir una taxonomía que refleje la terminología del ámbito futbolístico y que cumpla con las normas y directrices bibliotecológicas para construir tesauros.

CONSIDERACIONES FINALES

El panorama precedente muestra algunos de los últimos proyectos de organización temática de la información, desarrollados, la mayoría, con apoyo gubernamental de alto nivel. En cuanto a la metodología aplicada no siempre es visible, sin embargo, se puede observar que algunos de ellos son producto de la práctica como el *Herein Thesaurus*, pero con un gran estudio de las equivalencias lingüísticas entre idiomas. Se usan metodologías bibliométricas para analizar los

agrupamientos temáticos y el método comparativo para acercar los términos con las normas y los controles de autoridad. También se usa el método de la ruta crítica para desarrollar uno de los proyectos de mayor envergadura en el ámbito actual de la Bibliotecología.

Del examen de ésta pequeña muestra de proyectos, y en respuesta a las preguntas planteadas al principio de este trabajo, se puede decir que el proceso de trabajo difiere para cada investigación ya que la situación comunicativa que resuelven también es diferente. Otra observación que se puede extraer es que en cada práctica existe la misma concepción del tesoro. Por último la realización de tesauros como producto de investigación ofrece una mayor solidez en las propuestas para indizar contenidos y recuperar información.

OBRAS CONSULTADAS

Angel Santillan, Minerva del y Rosa Valganon, María Patricia. *Productividad y procesos técnicos : el caso de la DGB-UNAM*. –México : El autor, 1993. Tesis Licenciatura (Licenciado en Bibliotecología)-UNAM, Facultad de Filosofía y Letras. 194 p

Caminotti, María Luisa y Ana María Martínez “Fútbol, tesauros y taxonomías WEB: desafíos del control del vocabulario” p. 73-81 En *Información, Cultura y Sociedad*, No. 14 (2006).

Chan, Lois Mai y Marcia Lei Zeng “Ensuring interoperability among subject vocabularies and knowledge organization schemes: a methodological analysis” p. 1-5. En *IFLA Council And General Conference* (68: agosto 18-24, 2002: Glasgow).

Cisneros Rodríguez, Inés, García Doctor, Catalina y Lozano Jurado Isabel María (2000) *¿Sociedad de la información vs Sociedad del conocimiento? : La educación como mediadora*. (Consultada el 3 de septiembre de 2006).
<http://tecnologiaedu.us.es/edutec/paginas/43.html>

Escamilla González, Gloria *Manual de metodología y técnicas bibliográficas*. México: UNAM; Instituto de Investigaciones Bibliográficas, 1998. 161 p. (Instrumenta Bibliographica).

Tópicos de investigación en Bibliotecología y sobre la Información ...

European Commission. "Ontology-enabled e-Gov Service Configuration". Project IST-507232. (Consultada el 3 de enero de 2007)
<http://www.ontogov.com>

López Rico, Ma. del Pilar. *Comportamiento informétrico en la representación de contenidos en ciencias bibliotecológica y de la información relativo a América Latina*. México DF: La autora, 2006 (Tesis de Maestría) UNAM.

Moreiro González, José Antonio *El contenido de los documentos textuales: su análisis y representación mediante el lenguaje natural*. – Gijón: Trea, 2004. – 291 p. – (Biblioteconomía y administración cultural; 97).

University Of Portland State. "SGER: Accelerated Indexing in a Domain-Specific Digital Library". (Consultada el 4 de enero de 2007)
<http://www.digitalgovernment.org>

University Of Strathclyde. "High-Level Thesaurus". (Consultada el 4 de enero de 2007).
<http://hilt.cdlr.strath.ac.uk>

Vargas Suárez, Verónica Eugenia. *Análisis del lenguaje científico sobre agua*. México DF: La autora, 2007. (Tesis de Maestría) UNAM.

Zins, Chaim. Knowledge 2006: Map of Human Knowledge. (Consultada el 20 de noviembre de 2006).
<http://www.success.co.il/knowledge/about.html>