

Beneficios de la utilización de lenguajes controlados en el análisis y recuperación de información

MARÍA TEXIA IGLESIAS MATURANA

Profesora Titular de la Universidad Tecnológica
Metropolitana, Santiago, Chile.

Artículo

RESUMEN

Considerando la gran importancia de una ágil y adecuada transferencia de información en una sociedad en permanente cambio y marcada por la globalización y por un acelerado desarrollo científico y tecnológico, el papel de los especialistas en información cobra un cariz cada vez más preponderante. Dentro de este contexto, el diseño y desarrollo de buenos lenguajes documentales que ayuden a una mejor comunicación entre el autor y el usuario en las distintas áreas del conocimiento, parece una tarea de gran relevancia. Se discuten los beneficios que tiene la utilización de lenguajes controlados en el análisis y recuperación de información, versus el uso de un lenguaje libre.

ABSTRACT

Considering the major importance for quick and adequate information transfer within a society in permanent change and marked by globalization, and an accelerated scientific and technological development, the prospect of information professionals takes on an increasingly predominant role. A relevant task within this context would seemingly be, the design and development of sound documentary languages that help to improve the communication between author and user. The benefits of using controlled languages in information analysis and retrieval, versus the use of common language, are discussed.



Trabajo recibido
el 31 de octubre
de 1997

INTRODUCCIÓN

El propósito de este trabajo es entregar algunas consideraciones, basadas principalmente en la propia experiencia, tanto laboral como académica, y en las opiniones de diversos autores, sobre la conveniencia de usar lenguajes controlados en la indización y en la posterior búsqueda de información.

Hace un par de décadas, cuando algunos de nosotros nos iniciamos en el mundo de la información, el gran impacto era hablar de la **explosión de la información**, hoy en día la explosión de entonces nos parece un pequeño retumbar. Lo que en realidad, ahora ha exacerbado la situación ha sido la “Explosión

de la publicación”. **Publica o perecerás**, se dice en el mundo científico y académico. Esto significa que, muchas veces, una investigación, un trabajo o una experiencia dan lugar a una multitud de escritos, que, a veces, sólo se diferencian en el título. De manera que estamos literalmente ahogados en una avalancha de publicaciones.

Ahora bien, formamos parte de una sociedad marcada por la globalización y el cambio. El conocimiento, apoyado por la tecnología de las comunicaciones, surge, se reproduce y se difunde a una gran velocidad y por todo el mundo.

Junto a este panorama, tenemos los avances en la **tecnología de la información**, con sus supercarreteras, Internet, las bases de datos de texto completo, el hipertexto, los multimedios, la biblioteca virtual, etcétera. Para algunos, esta tecnología de la información es la panacea y la solución a todos los problemas. En lo personal, creo que estos avances tecnológicos nos han solucionado muchos problemas, pero debemos reconocer que nos han obligado a enfrentar otros.

En este escenario, pienso que el papel de los especialistas en información cobra un cariz cada vez más preponderante para lograr una ágil y adecuada transferencia del conocimiento. Dentro de este contexto, el diseño y desarrollo de **buenos lenguajes documentales** que ayuden a una mejor comunicación entre el autor y el usuario en las distintas áreas del conocimiento, parece una tarea de gran relevancia.

A fin de colocar el tesoro, el vocabulario controlado por excelencia, en su justa perspectiva, pensemos en algunos de los aspectos más convencionales de nuestro quehacer.

La práctica bibliotecaria tradicional se preocupó por arreglar los libros en un orden generalmente basado en un entendimiento filosófico del conocimiento. Los esquemas que regían este orden eran jerárquicos y enumerativos y el trabajo del indizador era ubicar el libro en el “casillero” correcto.

El deber del esquema de clasificación era proveer suficientes “casilleros” para la tarea de los indizadores, la cual requería una cierta dosis de flexibilidad para acomodarse con la expansión del conocimiento. Mientras el conocimiento se expandió de manera uniforme y relativamente lenta, los sistemas jerárquicos más o menos funcionaron.

Sin embargo, el extraordinario avance del conocimiento, como ha ocurrido, en especial, en las áreas de la electrónica y la biología, y el que nuevas temáticas se empezaran a conformar a partir de la conjunción de otras áreas, como la bioingeniería por ejemplo, hizo evidente la necesidad de un enfoque fundamental y multifacético para la organización del conocimiento.

Dice un autor, ya en 1965 “[...]debemos tratar de disociar la información de la literatura, no queremos libros, queremos información, y aunque ésta esté contenida en libros, debe observarse como un material diferente[...]tratarse diferente[...] debemos transformar la literatura en pedazos y reordenar sistemáticamente estos pedazos[...] a través de este método veremos que es posi-

ble alcanzar una exactitud casi matemática en la manipulación y coordinación de la información”. (Kaiser, 1965.)

El principal avance en relación con la clasificación enumerativa, se atribuye a Ranganathan, quien publicó en 1933 su Clasificación de Los dos Puntos (Colon Classification), de hecho, fue el primer esquema en utilizar, en realidad, el principio analítico-sintético.

Basados en este principio analítico-sintético, los esquemas se construyen de una manera inductiva a partir de ciertos conceptos fundamentales, los cuales pueden combinarse de diversas formas para conformar un compuesto sintetizado. Los conceptos fundamentales se logran mediante el análisis y el arreglo en facetas y categorías. Esto da al indizador una total libertad para construir notaciones relevantes para el documento que está describiendo; para construir, de alguna manera, sus propios “casilleros”.

INDIZACIÓN Y RECUPERACIÓN (INPUT & OUTPUT)

En 1965, Vickery describió sucintamente el proceso de indización, como una operación de tres etapas:

- 1) Revisión del texto para seleccionar un conjunto de palabras, frases o sentencias que colectivamente representen su contenido temático.
- 2) Tomar una decisión respecto a cuáles de estas descripciones temáticas vale la pena registrar como relevantes para los propósitos del sistema de recuperación.
- 3) Las descripciones temáticas relevantes se transfieren a un lenguaje descriptor normalizado en uso en el sistema.

Por otra parte, Wilfrid Lancaster (para quien la indización es un proceso de clasificación) describe el proceso de búsqueda, diciendo:

Al conducir una búsqueda en un sistema de recuperación de información para satisfacer una necesidad de información en particular, nosotros:

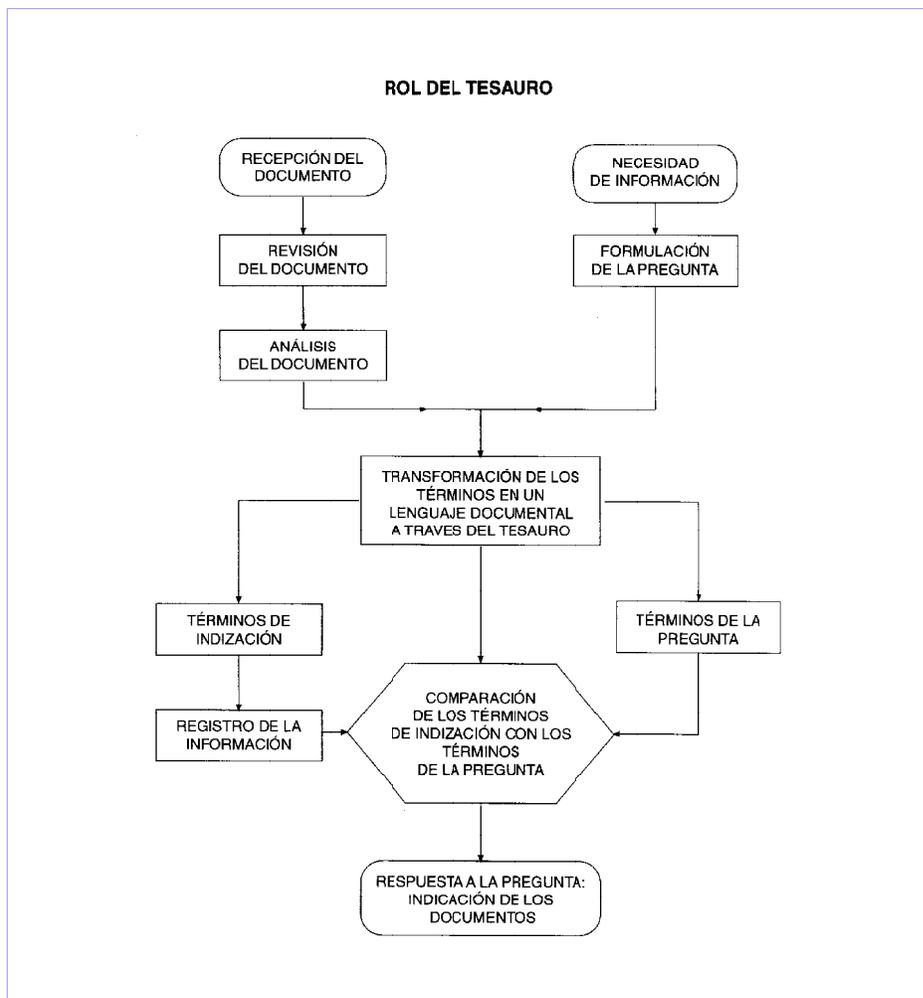
- 1) Decidimos qué clases (término de un lenguaje descriptor normalizado) es más probable que contengan los documentos relevantes para la necesidad de información.
- 2) Examinamos esas clases.
- 3) Recuperamos algunos o todos los documentos. (Lancaster, 1972.)

Así puestas las cosas, consideramos el lenguaje descriptor normalizado como el principal agente para establecer una coincidencia entre los vocabularios de indización y de búsqueda, al realizar ambas operaciones utilizando un mismo lenguaje.

Como dice Marie Thérèse Laureilhe, “[...]es un elemento esencial de la cadena que asegura la ligazón entre el documento y el usuario”. (Laureilhe, 1977.)

En otras palabras, el vocabulario controlado es el puente entre el lenguaje del autor, el lenguaje del usuario y el lenguaje del indizador u operador del sistema.

Figura 1
(Lancaster, 1978)



PRE-COORDINACIÓN VS. POST-COORDINACIÓN

La “indización coordinada” era considerada novedosa hasta hace algunos años, aunque de hecho la gente la ha estado utilizando desde hace mucho tiempo. Lo que sí surgió en la década del 50, fue la indización “post-coordinada”, en oposición a la indización “pre-coordinada.”

La indización pre-coordinada significa que la coordinación de conceptos diferentes, con el propósito de expresar una temática compleja, se realiza al momento de la indización, construyendo términos que representen la conjunción temática que se desea alcanzar, como aún lo podemos ver en los tradicionales catálogos en fichas.

En los sistemas post-coordinados, en cambio, la coordinación de conceptos separados se realiza al momento de la recuperación, como normalmente hacemos hoy al preparar nuestras estrategias booleanas y confrontarlas en nuestras búsquedas en bases de datos computarizadas. Y, aunque sin duda, estos sistemas post-coordinados solucionaron la gran dificultad que existía para expresar temas complejos (de hecho, formados por la conjunción de varias temáticas diferentes), también trajeron el gran ruido de las falsas coordinaciones.

¿POR QUÉ UN VOCABULARIO CONTROLADO?

Los sistemas post-coordinados, con la tecnología de la información disponible hoy en día, permiten realizar búsquedas en texto libre, obviando el proceso de indización.

Esto significa que no es indispensable el uso de un vocabulario controlado o “lenguaje descriptor normalizado” como lo llama Vickery, para realizar nuestras búsquedas en el mencionado texto libre, representado, generalmente, por un resumen. Aún más, podemos realizar esas búsquedas en lenguaje libre y en texto completo.

Sin embargo, que algo no sea indispensable, no significa que esté de más o sea descartable; muy por el contrario, ese algo puede ser bastante necesario, beneficioso y útil. Esto es lo que ocurre, en mi opinión, con los vocabularios controlados, utilizados como puente entre el autor y el usuario. (Ver *Figura 2*)

El rol principal de un vocabulario controlado en un sistema de recuperación de información, nos lo muestra este, muy conocido, diagrama de Lancaster. Permite la confrontación de los perfiles de preguntas y documentos, ambos expresados en el mismo lenguaje.

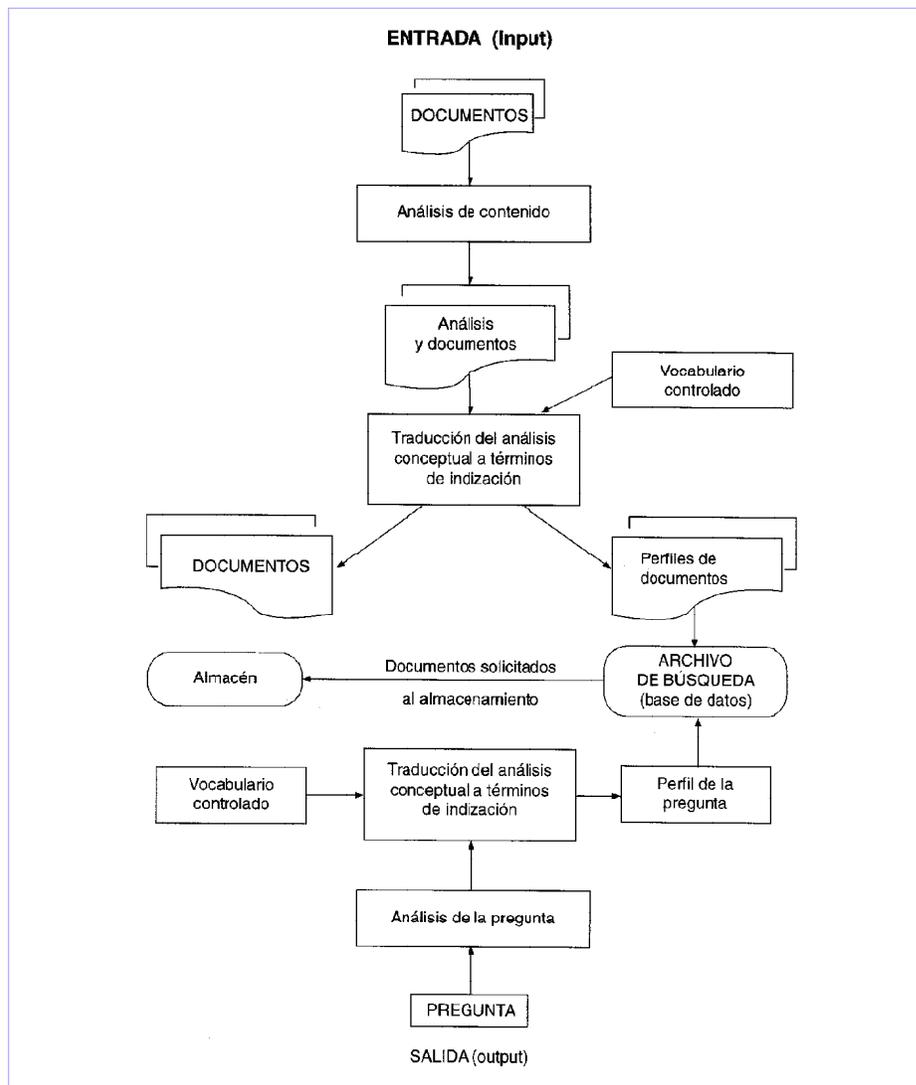
No obstante, aquí hay un punto importante que sería bueno recordar: el vocabulario controlado no influye, o por lo menos no debe influir, en el análisis conceptual de los documentos ni en el de las consultas. La etapa del análisis conceptual es aparte y previa a la etapa de traducción al lenguaje del sistema. Primero decidimos de qué trata un documento o una pregunta, y luego tratamos de traducir nuestro análisis conceptual a términos del lenguaje de indización.

Estas dos etapas, la de análisis y la de traducción al lenguaje, tienen diferentes efectos en el performance del sistema de recuperación de información.

Por ejemplo, luego de analizar conceptualmente la pregunta planteada a nuestro sistema, podemos tener muy claro y decidido que lo que nuestro usuario necesita

son documentos que traten sobre el *síndrome de las vacas locas*, que expresado en términos más científicos sería *encefalopatía espongiforme del bovino*.

Figura 2
El sistema de recuperación de información

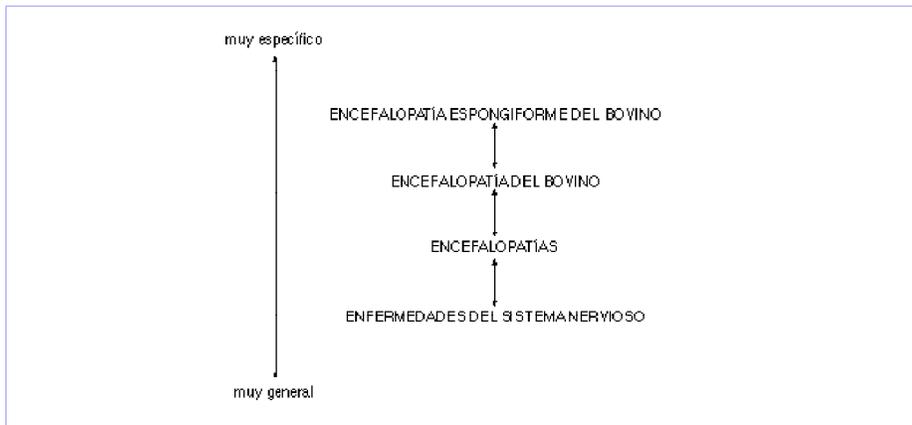


(Lancaster, 1978)

Éste es el resultado de nuestro análisis conceptual. Al ir a nuestro vocabulario puede que no encontremos el término y nos veamos obligados a expresar nuestra

búsqueda en términos menos precisos, más genéricos, como *encefalopatías del bovino*, por ejemplo; o aún más general, como *encefalopatías*, o aún peor, *enfermedades del sistema nervioso* (Ver Figura 3)

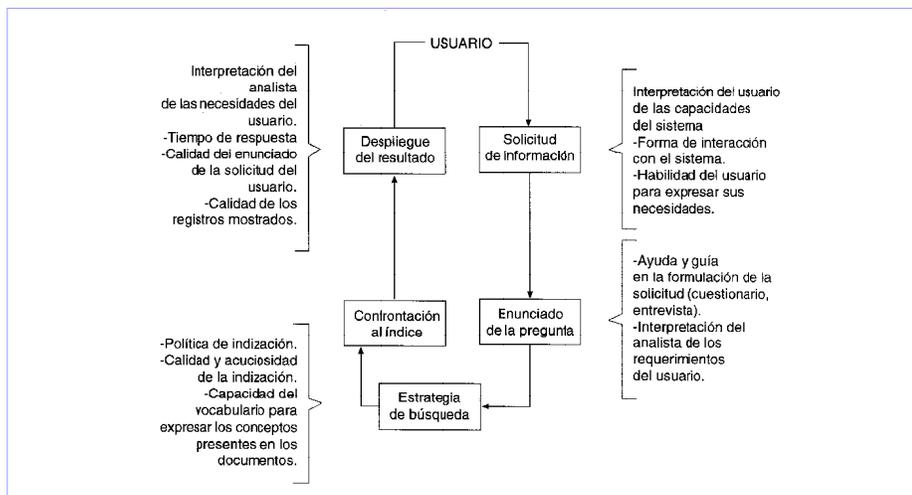
Figura 3
Especificidad del Vocabulario



Esta falta de especificidad del vocabulario afectará sin duda la precisión en el resultado de la búsqueda, aún cuando ciertamente favorecerá el *recall* de la misma.

Pero, observemos ahora otro diagrama. Éste nos muestra los diversos pasos de una operación de recuperación de información y, al lado de cada uno, los factores que más significativamente afectan el resultado. (Ver Figura 4)

Figura 4
Pasos en el proceso de recuperación y factores que afectan el éxito de cada paso



Como vemos, al observar el diagrama, el lenguaje de indización afecta el performance en dos de los principales puntos. Afecta la **estrategia de búsqueda**, al determinar cuán precisamente podemos describir los intereses de información de un usuario en particular, y afecta la **indización**, al establecer cuán precisamente puede el indizador describir el contenido temático de un documento. El uso o no uso de un vocabulario controlado es una parte muy importante en la operación de recuperación, pero sin duda que afecta el performance del sistema como un todo.

EL VOCABULARIO CONTROLADO COMO UN MODELO CONCEPTUAL

En este contexto y considerando la problemática del análisis y recuperación de información con un **enfoque sistémico**, podemos asociar un lenguaje controlado con el “modelo conceptual de un sistema”. “Un **modelo conceptual**”, dice Codina:

[...]expresa los elementos fundamentales de un sistema, haciendo abstracción de los detalles particulares de las diversas ocurrencias de los sistemas reales existentes. El objetivo de un modelo conceptual es proporcionar una representación del objeto o fenómeno que permita su comprensión global, como un todo, sin que tal comprensión quede dificultada por la confusión entre las cuestiones accidentales, que sólo se dan en algunas ocurrencias del modelo. (Codina, 1994.)

Los modelos conceptuales también facilitan la comunicación entre equipos de personas implicadas en tareas comunes, al eliminar la ambigüedad de ciertos aspectos del fenómeno considerado, y posibilitar el uso de un lenguaje común entre los miembros del equipo, evitando los malos entendidos, en especial, dados a aparecer, paradójicamente, en todos los asuntos relacionados con el desarrollo de sistemas de información.

Siguiendo a Walker, un sistema de información “es un sistema que mantiene alguna clase de registros sobre las características y la evolución de otro sistema [colección de documentos], el cual suele designarse como sistema del mundo real o como sistema objeto.” (Walker, 1991.)

“El sistema de información” agrega Codina, “no registra todas las variables o parámetros posibles del sistema del mundo real (o sistema objeto) que son virtualmente infinitos, sino que abstrae y selecciona ciertas características del sistema objeto de acuerdo con los propósitos del sistema de información y con los intereses de sus usuarios.” (Codina, 1994.)

Si dentro de un sistema de información, así considerado, tomamos, a su vez, el vocabulario controlado como un “modelo conceptual”, vemos que éste no registra todos los vocablos del lenguaje real (o sistema objeto) sino que selecciona aquellos que considera más relevantes, más específicos, menos ambiguos, de

acuerdo con los propósitos específicos de un sistema de información en particular y con los intereses de sus usuarios.

Así, entonces, el vocabulario controlado facilita la comunicación y la transferencia de información, al eliminar la ambigüedad de ciertos aspectos del fenómeno controlado (en este caso, el lenguaje natural y libre en que está escrito un documento), privilegiando los conceptos sobre las palabras. Ahora bien, aunque el modelo construido por el sistema de información (vocabulario controlado) debe obedecer a los propósitos particulares del sistema y a los intereses de sus usuarios, también debe ser fiel al sistema objeto (o lenguaje real) para ser eficiente. En consecuencia, es posible construir un lenguaje (visto como un modelo conceptual) que comunique al autor, el indizador y al usuario.

Este lenguaje de indización, en primer lugar, aporta un control sobre los sinónimos y los homógrafos, pero además, lo que es muy importante, permite realizar **búsquedas genéricas**. Al establecer una red de referencias cruzadas entre los términos, de alguna manera reúne las temáticas relacionadas, permitiendo realizar una búsqueda sobre una temática amplia.

Supongamos que al usuario le interesa el tema de los “esteroides”. En este caso, el vocabulario debe ser capaz de mostrar todos los términos relevantes que tengan que ver con “esteroides”, a partir de los TG (términos genéricos), los TE (términos específicos) y los TR (términos relacionados). Esta característica, concretamente de los tesauros, alivia el esfuerzo del usuario, evitándole el tratar de imaginar todos los posibles términos relacionados con “esteroides” bajo los que tendría que buscar, para realizar una búsqueda exhaustiva. Sin mencionar que ya le ha evitado todos los posibles sinónimos y cuasi-sinónimos a través de las referencias “USE”.

LA BÚSQUEDA EN TEXTO LIBRE

No es mi intención hacer una diatriba en contra de la búsqueda en texto libre como una forma de apoyar mi defensa del uso de vocabularios controlados. Sin embargo, permítanme transcribir lo que Reva Bash llama “Los siete pecados capitales de la búsqueda en texto completo”, sin duda sufriente de todos los problemas de la búsqueda en texto libre... y algo más. Sólo me detendré un momento en el “Pecado N° 2”: VERBOSIDAD.

Pecado N° 1: DUPLICIDAD. El texto completo en línea, rara vez es el mismo texto completo impreso.

Pecado N° 2: VERBOSIDAD. Tantas palabras... tan pocos descriptores.

Pecado N° 3: DIFICULTAD (Wimpiness). Tanto los usuarios como el sistema deben desarrollar nuevos músculos.

Pecado N° 4: IRRELEVANCIA. Las opciones de despliegue deben ayudar y no esconder.

Pecado N° 5: SUCIEDAD. Oh, esos datos sucios.

Pecado N° 6: HIPÉRBOLE. ¿Búsqueda global o “nunca mezcle, nunca se preocupe”?

Pecado N°7: OFUSCACIÓN. “Finalmente logré reunirlo todo... pero olvide dónde lo puse”.

La verbosidad

Numerosos estudios describen la negociación entre *recall* y precisión en las búsquedas en texto completo. En términos sencillos, poner más palabras en la “alberca de búsqueda” significa más aciertos en sus términos de búsqueda, pero una menor proporción de aciertos relevantes. Esto porque, en general, las bases de datos de texto completo carecen de alguna suerte de indización detallada o clasificación jerárquica que facilite la búsqueda.

El inmenso volumen de información en los archivos de texto completo, combinado con la natural “caja negra” de las búsquedas en línea, hace que la falta de puntos de acceso controlado sea un gran pecado de omisión.

Reva Bash entrega un pequeño recuento de los trucos que es necesario usar para compensar esta “verbosidad” y falta de vocabulario controlado.

- 1) Utilice jerga, sinónimos y antónimos. Construya su propia fuente de conceptos que utiliza frecuentemente.
- 2) Realice una **truncación** que permita la incorporación de los plurales y otras variaciones y terminaciones verbales alternativas.
- 3) Use operadores de proximidad y otros dispositivos que mejoren la precisión y el análisis de los registros recuperados en busca de relevancia.
- 4) Restrinja los términos de búsqueda al mismo tema o párrafo.
- 5) Realice su búsqueda, primero, en una base de datos bibliográficos para aprovechar las **ventajas de su vocabulario controlado**, luego trasládese al texto completo para recuperar artículos específicos. De esta forma, usted estará usando la base de texto completo como una forma de obtener documentos, en lugar de una herramienta de búsqueda. (Bash, 1989.)

EL TESAURO COMO VOCABULARIO CONTROLADO

He procurado mostrar los aspectos positivos de utilizar un vocabulario controlado en el análisis y recuperación de información. Sin embargo, si queremos que tal lenguaje sea eficiente y realmente mejore la recuperación de información, es necesario que cuente con algunos **atributos** como los siguientes:

- 1) Debe contar con un **vocabulario** y una **sintaxis** comprensibles para el que busca. Digo para el que busca, porque hoy en día muchas veces esa persona es el **usuario final**. El indizador, por razones obvias, conoce bien el lenguaje de indización que usa. Si los usuarios del sistema de información son público general, o por lo menos no particularmente especializado, el lenguaje de indiza-

ción debe aproximarse en lo posible al lenguaje natural. En una unidad de información, en extremo, especializada, el vocabulario debe reflejar las expresiones que son propias de los especialistas. En el primer caso, es posible que usemos *enfermedad de las vacas locas*, y en el segundo caso incluiríamos *encefalitis espongiiforme del bovino*.

- 2) Debe ser capaz de aceptar sus propias modificaciones, en la medida que cambia el lenguaje de los documentos o de los usuarios.
- 3) Si el archivo de búsqueda consiste en una base computarizada de datos bibliográficos, implica que el formato, la sintaxis y el vocabulario del lenguaje de indización deben ser comprensibles para la máquina y para el hombre.
- 4) Como lista de autoridad debe, además, tener las siguientes características:
 - a. **Control de sinónimos** para evitar que documentos iguales se indiquen de diferente forma, lo que ocasiona que se dispersen y se dificulte su recuperación.
 - b. **Control de homógrafos** para evitar la confusión entre palabras que se escriben igual y tienen diferente significado.
 - c. **Genéricamente estructurado**, a fin de que muestre las relaciones jerárquicas entre los términos, lo que facilita, al momento de la indización, encontrar los términos más generales o más específicos según se necesite, y permite la búsqueda genérica al momento de la recuperación.
 - d. Con un **adecuado nivel de pre-coordinación** en sus términos, que evite, o por lo menos aminore, las falsas coordinaciones tan frecuentes en los sistemas post-coordinados.
 - e. Con las **necesarias notas de alcance**, vale decir, que tenga notas aclaratorias, dentro del contexto de ese vocabulario en particular, respecto a algunos términos que así lo ameriten para su mejor uso y comprensión.

Sin duda, un vocabulario controlado que suele reunir estos atributos y características es el **tesauro**.

El tesauro

Dentro del ámbito de la Recuperación de Información, según Vickery, la primera persona que usó por escrito la palabra “tesauro” en su aplicación común de hoy, fue Helen Brownsen en 1957 en una Conferencia sobre Clasificación realizada en Dorking, Inglaterra. Brownsen dijo:

[...]el problema [de la recuperación de información], al menos como algunos investigadores lo ven, es transformar los conceptos y sus relaciones como aparecen expresadas en el documento a otro lenguaje más reglamentado, con los sinónimos controlados y las estructuras sintácticas simplificadas[...]Algunos investigadores piensan que la mejor respuesta o solución[...]puede ser la aplicación de un *tesauro mecanizado* basado en redes o significados relacionados. (Vickery, 1960.)

En nuestro contexto actual, y coincidente con lo pensado por Helen Brown-sen, un tesoro es la herramienta utilizada para normalizar la terminología al indizar documentos con vistas a su posterior recuperación, y, también es la herramienta que se utiliza para indizar las preguntas que llegan a un sistema.

El hecho de utilizar la misma herramienta en ambos procesos, permite lograr una coincidencia entre los documentos (representados por sus términos de indización o **perfil** del documento) y las consultas (representadas por sus términos de indización o **perfil** de la pregunta). (Schultz, 1978.)

Un documento, entonces, será recuperado cuando se produzca una **coincidencia** total o parcial entre el perfil de la pregunta y el perfil del o los documentos almacenados en el sistema. (Iglesias, 1992.)

La experiencia de construir un tesoro

La construcción de un tesoro, como muchas otras actividades, puede enfocarse teórica o empíricamente. Quienes prefieren el método empírico, examinan el lenguaje que realmente utilizan los autores (seleccionados al azar) en el campo temático que les preocupa, y realizan muchas de sus decisiones por estadísticas, decisiones referidas a si incluir o no un término, o elegir entre sinónimos cuál será aceptado y cuáles se ingresarán con referencia de USE.

En contraste, quienes utilizan el enfoque teórico en la construcción de un tesoro suelen comenzar un poco como lo hizo Melvin Dewey, tratando de pensar “lógicamente” acerca de cómo el universo debe ser dividido y expresado. Generalmente este pensamiento se basa en un examen jerárquico del conocimiento más que en un examen del uso del lenguaje.

Quizá, un tesoro concebido con un enfoque teórico será menos eficiente tanto en la indización como en la búsqueda de información, aunque más estable en el tiempo. Pero, no es la estabilidad lo que más se busca en un tesoro, ya que por su propia naturaleza se espera que sea una herramienta viva y dinámica.

Con Catalina Naumis de México, Ana Cecilia Osuna de Venezuela y Lucy Espinoza de Colombia, desde hace cerca de un año se ha estado aunando esfuerzos en un proyecto interinstitucional, bajo el alero del CUIB y de INFOBILA, para la construcción de un Tesoro Latinoamericano en Ciencias Bibliotecológicas y Documentales.

Este tesoro, cuya versión preliminar se estará poniendo a prueba durante los próximos meses, es una conjunción del trabajo de personas diferentes que realizaron su labor en distintos países y con diversos propósitos.

En el caso de Chile y Colombia, inicialmente, fue el resultado de trabajos de titulación, después, desarrollados y ampliados por las profesoras guías; por ende, el propósito primario de estos tesoros fue de orden académico, por lo que puede decirse que su enfoque de construcción es una mezcla entre el empírico y el teó-

rico. En el caso de Venezuela y México, en cambio, los aportes fueron el resultado de un trabajo experimental, donde la terminología fue surgiendo a medida que se indizaron cantidades importantes de documentos.

Durante el desarrollo de este trabajo y al tratar de fundir en un producto el esfuerzo de todas, se comprobó cuán diferente puede ser el uso del idioma español en países latinoamericanos, aún tratándose de una misma área del conocimiento. Naturalmente, no se podía dejar de pensar cuán dificultosa podría resultar la tarea de preparar estrategias de búsqueda para recuperar documentos en texto libre, en una base de datos construida con el aporte de todos los países latinoamericanos sobre la base de resúmenes. Digamos concretamente INFOBILA.

Sin embargo, y más allá de la utilidad que este tesoro pueda tener para la indicación y recuperación de documentos de la especialidad, subyace un objetivo muy importante, que es lograr una normalización en el lenguaje utilizado por los profesionales de información en América Latina.

Una disciplina, además de sus fundamentos teóricos y filosóficos, necesita de una **terminología** inequívoca entre los pares, para avanzar en sus pretensiones científicas.

Finalmente, aun desde un punto de vista práctico, si el objetivo se logra, sin duda, será un gran aporte para el intercambio de experiencias profesionales, académicas y de investigación en la región.

BIBLIOGRAFÍA

- Borko, Harold and Bernier, Charles (1978). "Thesauri". *Indexing Concepts and Methods*. New York, Academic Press, pp. 93-111.
- Browne, Glenda (1992). "Scope notes for LISA subject headings". *Online Review*. 16(1) : 3-15, 1992.
- Fischer, Dietrich H. (1991). "Consistency rules and triggers for Thesauri". *International Classification*. 18 (4) :212-233.
- Gilchrist, Alan (1971). *The Thesaurus in Retrieval*. London, ASLIB.
- Iglesias Maturana, María Texia (1992). *Análisis y Recuperación de Información Documental*. Santiago, Ediciones Instituto Profesional de Santiago.
- Kaiser, J. (1966). "Systematic Indexing" Olding, R.K. *Reading in Library Cataloguing*. London, Crosby Lockwood.

- Lancaster, F. Wilfrid (1978). *Information Retrieval Systems*. 2nd. De. New York, Wiley.
- – (1972). *Vocabulary Control for Information Retrieval*. Washington, IRP.
- Laureilhe, Marie-Thérèse (1977). *Thesaurus, son rolw, sa structure, son elaboration*. Paris, E.N.S.B.
- López Morales, Carmen Yasmina (1995). “Infraestructura y servicios de información en la biblioteca virtual”. *Ciencias de la Información*. 26(3) : 115-123, septiembre.
- Schultz, Claire K. (1978). *Thesaurus of Informacion Science Terminology*. Metuchen, N.J. The Scarecrow Press. pp. 243-288.
- UNAM/CUIB (1996). *Tesauro Latinoamericano en Ciencias Bibliotecológicas, Documentales y de Información* (documento preliminar). México.
- Vickery, B.C. (1960). “Thesaurus - a new word in documentation”. *Journal of Documentation*. 16(4) : 181-189.
- Walker, D.W. (1991). *Sistemas de información basados en ordenador*. Barcelona, Marcombo.