

Servicios a usuarios:

La recuperación de la información y los lenguajes documentales

Por Zaida Sequeira, M. A.

Introducción

En este último cuarto de siglo XX lo que fue apenas necesario solo hace diez años, hoy es imperioso, tratándose de la recuperación de la información. La inmensa masa de información que hoy circula en el mundo, rompe toda posibilidad humana, individual o colectiva de abarcarla, de almacenarla y lo más importante, de darla a conocer.

La forma tradicional de buscar materiales para sacar información va quedando atrás, corta ante un cúmulo tan grande de bibliografía, general o especializada que hoy se publica.

Las bibliotecas que no utilizan técnicas modernas para entresacar, de toda esta información la que le conviene al lector, y siguen manejando documentos con la lentitud con que se manejaron hace un siglo, están limitado el servicio, que deben ser dinámico y acorde con los tiempos que se viven, a un servicio que ya no cumple con las necesidades para las cuales fue creado.

Es por esto que las bibliotecas, en muchos casos se han visto prácticamente desplazadas por los llamados centros de documentación, porque estos centros agilizan, activan una información que quizá está durmiendo, en los estantes y la entregan al lector de hoy siempre urgido de información, pero necesitando esa información lo más completa posible, también hoy.

Para cumplir estas demandas del "utilizador" de la información, se ha creado una serie de lenguajes orientados hacia la mejor comprensión de los documentos, y una serie de técnicas manuales, mecanizadas o computarizadas, que permiten la recuperación de la información que contienen esos documentos.

En este artículo se señalan en forma general las técnicas automatizadas y algunos de los lenguajes que se han creado para manejarlas. En un próximo artículo se tratará la recuperación manual, basada en estos conceptos preliminares.

Tipología de la consulta

Todo sistema de información documental tiene como fin explotar la información y nos reduce esos objetivos a tres estados que son:

- a.- Búsqueda bibliográfica retrospectiva
- b.- Servicio de difusión selectiva de información
- c.- Elaboración de boletines e índices.

La consulta puede realizarse:

- a.- Períodos fijos:

Como es la difusión selectiva de información en la que el utilizador o lector expone sus temas de interés y periódicamente el sistema ofrece referencia documentales.

b.- Forma aleatoria:

Un ejemplo es la búsqueda retrospectiva que son referencias que responden a la preocupación del utilizador y que se encuentran en todo el conjunto del fondo del documento.

Proceso de recuperación documental

La recuperación de información se resume así:

- a.- El usuario hace la demanda documental dando las características de información que necesita.
- b.- Los documentalistas o bibliotecarios resumen la demanda siguiendo ciertas técnicas.
- c.- El proceso de recuperación selecciona las referencias documentales que se solicitan.
- d.- El resultado son listas de referencias, que nos permiten saber:

- 1. Si existen documentos sobre el tema que se les pide y cuáles son.
- 2. Dónde se han publicado y cómo obtenerse, si son originales o traducciones.

En los distintos tipos de búsqueda los pasos que siguen los documentalistas y los usuarios en cuanto a la demanda, registro y presentación de referencias son muy parecidas.

Como se realiza la recuperación

La recuperación de la información consiste en seleccionar referencias documentales que permiten localizar y obtener la información primaria. Todo artículo que pasa al fondo documental entrará en forma de referencias siguiendo ciertos criterios establecidos.

Todo artículo si se desea recuperar se obtendrá en forma de referencias y se usarán los mismos criterios que se usaron para realizar el resumen; a esos criterios se les llama campos de recuperación

Los datos que se usan son los siguientes:

- Nombre del autor
- Fecha de publicación
- Entidad que lo publica
- Palabras claves del thesaurus
- Palabras claves extraídas del título
- Palabras claves extraídas del texto
- Términos del lenguaje libre

Evolución de la búsqueda documental

El proceso se realiza mediante ordenadores pero aún así el resultado no llega rápido al usuario; ya que la introducción de la consulta en el sistema requiere mucho tiempo porque también hay que compilar el programa, enviar la respuesta al solicitante y el resultado deberá analizarse para saber si sirvió la estrategia de búsqueda.

Si existieran problemas en cuanto a la recuperación, se pueden resolver por medios manuales, si el documentalista usara los índices cruzados que nos da el ordenador solo que es un trabajo

largo y difícil ya que son muchos los criterios que se seleccionan.

Proceso de consulta interactiva en tiempo real

Es el proceso de recuperación que tiene las siguientes características:

- a.- El fondo documental está almacenado en la memoria de un ordenador.
- b.- Existe un programa que trabaja con el ordenador en forma de diálogos.
- c.- El diálogo produce una interacción entre el documentalista y el ordenador; y al mismo tiempo en que se efectúa la consulta se orienta a seguir con la estrategia de la búsqueda.

La consulta interactiva se puede realizar a centenas de kilómetros del ordenador central y se le llama redes de tele documentación.

En el archivo directo las referencias quedan ordenadas con secuencia, tal y como entran en el fondo documental, y en el archivo inverso, están ordenados por palabras claves con los documentos indizados por dichas palabras.

En un sistema documental automático los archivos son de referencias documentales completas (autor, título, pie de imprenta, resumen de cada documento) y estos datos son los que se introducen en el ordenador ya que luego se usan para recuperar los documentos que más tarde solicita el documentalista.

El documentalista tiene a su disposición una pantalla de terminal, un teclado correspondiente de signos alfabéticos y operativos, el teléfono y una impresora.

Los pasos a seguir en una recuperación de información son los siguientes:

- a. Se recibe la demanda y se observa si se puede dar una respuesta lógica.
- b. Se busca en el thesaurus la palabra clave para empezar la búsqueda; si estuviese incluido en el ordenador el documentalista teclea en la pantalla para verificar si existe.
- c. Luego se dirige al archivo inverso y revisa en la pantalla, para ver cuántos documentos aparecen en esa memoria, de cada descriptor.
- d. Cuando la respuesta del número de documentos se considera correcta, el documentalista obtiene la posibilidad de que aparezcan en la pantalla las referencias bibliográficas con el resumen incluido.
- e. Se seleccionan los documentos cuando la lista es muy exhaustiva.
- f. Se recombinan los documentos.
- g. Una vez que el documentalista comprueba por medio de la pantalla que las referencias documentales obtenidas responden a la demanda, entonces el resultado queda listo en el ordenador central.

Proceso de consulta diferida

Los documentos se encuentran memorizados en el ordenador y se dispone de un programa que permite la recuperación de los documentos. Existen ciertas variantes que son:

- a. No permite forma dialogada de consulta, sino que obliga a introducir las demandas en forma de paquete en el ordenador.
- b. Los medios de consulta no son directos y las demandas se registran previamente.
- c. No permite cambiar la estrategia de la búsqueda. Porque el documentalista debe esperar a que le llegue la respuesta a la demanda; para saber si fue bien formulada.
- d. La consulta se efectúa mediante las tarjetas perforadas, previamente codificadas a partir de fichas muestras de interrogación en donde están inscritos los diferentes niveles lógicos.
- e. Las demandas están reunidas y codificadas en tarjetas de consulta, junto a los descriptores.

Lenguajes documentales

Estructura jerárquica: Por medio de la estructura jerárquica se establece una relación entre términos generales o específicos para que la información pueda encontrarse ya sea recurriendo a unos o a los otros.

Estructura asociativa: La estructura asociativa es la que permite establecer relaciones entre conceptos opuestos o cercanos.

Clasificaciones enciclopédicas: Un grupo de trabajo de la ONU examinó las varias clasificaciones que compiten en el conocimiento mundial, como normas para la categorización por material. Como la clasificación Decimal Dewey, la lista de Encabezamientos de Materia de la Biblioteca del Congreso, la Clasificación Decimal Universal, el grupo concluyó que ninguna da respuesta exacta a materias muy especializadas/.

Apuntan que falta alguna adecuación total o local en lo que atañe al contenido y a la estructura de las divisiones de la Clasificación Decimal Universal, estimándose que se requieren más estudio y experimentos para medir el valor potencial de la CDU en su estado presente, como lista mundial única de encabezamientos de materia para categorizaciones generales o indización superficial de documentos. Los experimentos realizados para mostrar que las clasificaciones enciclopédicas pueden ser utilizadas también como lenguaje de descripción para indización en profundidad y recuperación mecanizada.

Existe una gran distinción entre clasificaciones con rígidos encasillamientos, por una parte y lenguajes descriptores, absolutamente flexibles por otra.

Clasificaciones especializadas. Este tipo de clasificación es usada por dos grupos de personas, a saber:

- a. Aquellos cuyos intereses se concentran en un determinado asunto y afirman con razón o sin ella que la Clasificación Decimal Universal no se adapta bien o no es lo bastante detallado para sus especiales necesidades.
- b.- Quienes no pueden encontrar entre los existentes ningún sistema que se adapte a sus necesidades y en consecuencia se ven obligados a concebir un sistema de clasificación propia.

Ambos grupos deben tener en cuenta dos condiciones que ha de tener el sistema para ser practicable:

- a. Asignarse un lugar preciso a cada tipo de tema que pueda aparecer o preveer su posterior inclusión.
- b. Estos lugares o casilleros deben rotularse de tal manera que se excluyan entre sí.

Clasificaciones facetadas: La esencia de la clasificación por facetas según Foskett consiste en que persigue reflejar la diversidad y la multiplicidad de las relaciones entre los fenómenos que se dan en el mundo real.

Ella no persigue dar una lista de materias ordenadas por un sistema científico aristotélico a base género-especie, sino que analiza los fenómenos que ocurren en la naturaleza, ordenados en series homogéneas de categorías de cosas, procedimientos, etc..

Lenguajes documentales: El lenguaje documental es un lenguaje que puede estar basado o no en el lenguaje natural; es estereotipado normalizado y lo representa un término, un código o un símbolo.

Es el que utiliza para representar el contenido de un documento en sus elementos fundamentales. Es el que nos da una traducción del lenguaje natural, con la multiplicidad de sentidos de sus palabras, al vocabulario documentario que es preciso, controlado, normalizado,

lógico y asociativo y cuya comunicación no la obtenemos en forma inmediata, directa, sino a través de listas de conceptos o índices diccionarios.

Tesaurus: Se entiende como un conjunto definido y controlado de términos relacionados entre sí que transmiten para recuperar información en un ámbito dado.

Los tesauros se limitan a la elaboración de unas listas de términos normalizados constituidos en un lenguaje documentario.

Cada artículo impreso como consecuencia del pertinente examen por las palabras claves o por los descriptores del "tesaurus" que le sean aplicables. A tal fin procede a indagar la palabra o palabras que lo definen y traducirlas a las que se han elegido para representarlas en el tesauro.

Hasta ahora los tesauros vienen elaborándose con miras a su aplicación a campos especializados; ello no obstante, se espera poder elaborar algún día un tesauro universal que incluya todas las especialidades y materias.

Vikery: Propone una combinación aún más compleja que simboliza así: sustancia, producto, organismo (P); parte, órgano, estructura (O); constitución (C); propiedad y medida (Q); objeto de la acción, materias primas (R); acción, operación, procedimientos, comportamiento (C); agente, espacio y tiempo (ST).

Este lenguaje documental tiene una estructura jerárquica y clasificación facetada.

Dewey: De todas las clasificaciones en uso actualmente, la del bibliotecario norteamericano Melvin Dewey, la Clasificación Decimal Dewey, es posiblemente la que ha alcanzado mayor difusión, el sistema decimal distribuye la totalidad de los conocimientos en 10 clases a cada una de las cuales asigna un número de 0 a 9, estas clases que forman lo que podríamos llamar clasificación de primer grado se subdividen a su vez en otros, los que forman la clasificación de segundo grado y así sucesivamente, comprendiendo materias cada vez más específicas, hasta el infinito.

Este sistema es de estructura jerárquica y clasificación enciclopédica.

Bliss: Otro sistema de clasificación igual al de la Biblioteca del Congreso, que sigue un sistema análogo, que permite ser de un considerable interés, y que apareció bajo el título de "A System of Bibliographic Classification".

Bliss ha hecho un profundo estudio de un sistema que, al parecer se basa sobre relaciones más perennes entre diversas materias.

Este sistema posee una estructura jerárquica y una clasificación enciclopédica.

Patentes: Los objetivos de las colecciones de patentes son dos:

- a.- Facilitar con rapidez de una herramienta de trabajo a la investigación, ya que la lectura y el estudio de las patentes es una fuente de fértiles ideas y de fecundas iniciativas.
- b.- El desarrollo de la producción nueva.

Las patentes tienen por objeto primordial promover el progreso de la ciencia y de la técnica, garantizando a los autores e inventores el derecho exclusivo de explotar económicamente sus invenciones, aunque por un tiempo limitado. Es importante saber que las patentes se otorgan no solo en beneficio del autor del invento, sino también del público. Los países que producen mayor número de patentes son: Estados Unidos, Inglaterra, Alemania, Francia, Suiza y Japón.

De la perfección con que se aplique cualquier clasificación a la organización de un centro de consulta de patentes y el acierto en la clasificación elegida a este fin depende en su mayor medida, que este rinda el objeto que se persigue.

Recomiendan los diversos autores al tratar esta materia, que cada industria debe mantener y utilizar los boletines internacionales de información sobre patentes como complemento; recomendación que ha sido acatada por los técnicos de la mayoría de los países desarrollados.

Este sistema tiene una estructura asociativa y una clasificación facetada.

Ranganathan: Bibliotecario indú, creó el sistema que lleva su nombre y también el de clasificación por facetas o colonada, este sistema consiste en clasificar los asuntos no solo por la relación de género a especie, sino también por las existentes entre una cosa y sus partes, sus materias constitutivas, sus propiedades, los tratamientos a que pueden estar sometidas, el espacio y el tiempo.

Este sistema posee una estructura jerárquica y clasificación facetada.

Alfabético de Materias: Las listas normalizadas de encabezamientos de materias se destinan solo a las categorizaciones generales de documentos o indización superficial por oposición a la forma más profunda de análisis de contenido que se efectúan con las listas de descriptores y thesaurus especializados. Se estima que una lista normalizada de encabezamientos de materias generales puede ser útil cuando se trata de localizar, transferir grandes bloques de información, más que documentos y datos específicos.

Este sistema es de estructura asociativa y clasificación enciclopédica.

Uniterm: El sistema Uniterm fue concebido y realizado desde un principio como una técnica manual (1953), después esta técnica ha pasado a ser variada.

Este sistema parte de la elaboración de un "thesaurus" o lista de palabras claves de materias científicamente seleccionadas, con reglas e instrucciones que permitan relacionar los diversos temas en grado de profundidad.

Las palabras claves se toman del thesaurus a presencia de un examen exhaustivo de cada artículo, informe, relatorio, tesis, etc..

La palabra uniterm puede traducirse por palabra específica de materia. La manipulación del sistema es también sencilla, ya que solo requiere espigar en el documento las palabras claves que revelan su contenido y deben servir para su recuperación de acuerdo siempre con el thesaurus empleado.

Este sistema posee estructura asociativa y es un thesaurus.

Descriptores: La razón de ser lenguaje descriptor se basa en la necesidad de reducir la inmensa variedad de palabras utilizadas en los documentos y sustituirlos por un número más reducido y normalizado. Afirma Vikery, con acierto, que la introducción en un lenguaje de descriptores de relaciones jerárquicas y de coordinación permite reducir la dispersión de los términos conexos y encadenarlos semánticamente al objeto de lograr un mejor factor de respuesta.

En resumen: la lógica matemática nos ayudará a la organización de los descriptores y a la puesta en servicio de los selectores, la lingüística puede facilitar la selección de los descriptores.

Otros Lenguajes Documentales: Fue establecida originalmente por el Instituto Internacional de Bibliografía de Bruselas. Hoy se haya continuada y actualizada por la Federación Internacional de Documentación con asiento en la Haya.

La CDU es una ampliación y revisión del sistema de Dewey, pero conservando la estructura y características de este último. Ambos difieren en el hecho de que en el sistema Universal se ofrece un mayor desarrollo en las divisiones en la variante de algunas rúbricas especiales, además de la adopción de diversos signos auxiliares y las llamadas determinantes. En una palabra "su función pretende-como dice Penna, ser más específica, más bibliográfica.

Undex: La biblioteca Dag Hammarskjold, de la sede de las Naciones Unidas posee un sistema para almacenar los textos de amplios resúmenes y extractos de sus documentos, resoluciones de los órganos principales, etc. los que se aplican bajo el nombre general de UNDEX.

La información se extrae, selecciona y compila mecánicamente, en español, francés, inglés y ruso a partir de anotaciones preparadas únicamente en inglés, almacenados en computadora juntamente con las listas de términos en cuatro idiomas. Los términos consisten en grupos de palabras en mayúsculas, que describen el tema principal del documento y cuando resulta necesario palabras en minúsculas, que se refieren a detalles del tema principal de documento. Los Textos se conservan tanto en la forma original como en microfichas.

Pris: Muy recientemente se está ensayando un subsistema experimental para los proyectos de asistencia y cooperación técnicas titulado Project Reports Information Subsystem (PRIS). El sistema consistirá en listados de computadora (bibliográficos, de extractos, índices por países, índices por materias, etc.) relacionados con un sistema de reproducción de los proyectos en microfichas.

Sistema de documentación de la FAO

Una palabras claves a descriptores para formar frases de indización que describen los documentos. La FAO, posee un thesaurus de más de 8 mil descriptores y palabras, sinopsis del contenido del documento.

Sistema de la UNESCO

La UNESCO proyecta un servicio de documentación en computadora basado en el sistema de la OIT para la indización y quizás para la recuperación.

Conclusiones

El bibliotecario actual debe conocer lo nuevo que en el manejo de la información se está utilizando en el mundo. El bibliotecario de hoy no puede seguir pensando solo en su biblioteca de escuela, de colegio o de la universidad. Debe trascender las fronteras locales y fijar su atención en esfuerzos muy valiosos que se están realizando para lograr "sobrevivir" a la avalancha de información que nos inunda. No quiere decir esto que debemos pensar en automatizar nuestras bibliotecas como única alternativa. Eso sabemos lo difícil que es -pero no imposible- en nuestros países, tan pobres, tan retrasados. Sin embargo nuestra condición de subdesarrollados no puede ser una eterna excusa para no remozar nuestras técnicas, inclusive nuestra visión de lo que debe ser una biblioteca actualmente: debemos conocer qué se está ensayando en otros lados. Qué combinación manual creativa puede ayudarnos a dar cada día un mejor servicio, -razón de ser de todo bibliotecario-, y estar preparados para un cambio de siglo que se avecina, al cual viene precediéndolo la información, que no se detiene ni se estanca.