

**Ponencia**

**1**

# **PLANEACIÓN, DISEÑO Y DESARROLLO DE BASES DE DATOS BIBLIOGRÁFICAS**

Lda. María Ángela Torres Verdugo, Coordinadora del IRESIE (Índice de Revistas de Educación Superior e Investigación Educativa), Instituto de Investigaciones sobre la Universidad y la Educación, Universidad Nacional Autónoma de México. Miembro del Consejo Consultivo de REDMEX, México.



# **PLANEACION, DISEÑO Y DESARROLLO DE BASES DE DATOS BIBLIOGRAFICAS**

Lda. María Ángela Torres Verdugo, Coordinadora del IRESIE (Índice de Revistas de Educación Superior e Investigación Educativa), Instituto de Investigaciones sobre la Universidad y la Educación, Universidad Nacional Autónoma de México. Miembro del Consejo Consultivo de REDMEX, México.

## **INTRODUCCIÓN**

En la segunda mitad del siglo XX se produjo un incremento importante en el número de instituciones y personas que se dedican de manera profesional a la investigación en todos los campos del conocimiento, como consecuencia, el número de publicaciones: libros, artículos de revistas, ponencias, informes, materiales didácticos etc. día con día se multiplican. A su vez se hizo necesario el desarrollo de mejores herramientas para sistematizar la información y hacerla más accesible.

Para que la información se conserve y se constituya en fuente documental accesible en forma precisa y oportuna para diversos usuarios, pasa por una serie de procesos como: clasificación, sistematización, organización y almacenamiento, los cuales, por mucho tiempo, se llevaron a cabo en bibliotecas y centros de documentación en forma manual, pero a partir de la segunda mitad del siglo XX, la aplicación de la computadora al campo de la documentación propició cambios sustanciales.

Surgieron en EE. UU. sistemas de información como el Medline para sistematizar la información del área médica, y el ERIC (Educational Resources Information Center) para el campo de la educación. En Europa se crearon el EUDISED, que cubre la producción sobre educación de varios países de Europa, el Boletín Internacional de Bibliografía sobre Educación (BIBE), publicado en España con apoyo de la UNESCO, entre otros. En América Latina se instituyó REDUC: que integra el conocimiento generado en educación de varios países de la región.

En México, la creación de bases de datos bibliográficas se inició en la década de los setenta, tuvo su auge en los ochenta y noventa, principalmente en el ámbito de la Universidad Nacional Autónoma de México (UNAM). Se generaron, entre otras, bases de datos como:

LIBRUNAM, que contiene los registros de los libros existentes en sus 142 bibliotecas, SERIUNAM, catálogo colectivo que contiene la información de 53 000 títulos de publicaciones seriadas científicas y técnicas existentes en la UNAM y otras bibliotecas importantes de México, CLASE y PERIODICA, bases de datos que contienen más de 550 000 registros de artículos, ensayos, reseñas de libros, revisiones bibliográficas, entrevistas etc. publicadas en más de 3000 títulos de revistas de América Latina y el Caribe especializadas en Ciencias Sociales y Humanidades y Ciencia y Tecnología respectivamente, el IRESIE que contiene datos referenciales y el resumen de más de 55 000 artículos de revistas científicas y técnicas especializadas en educación que proceden de México y demás países Iberoamericanos.

### **DEFINICIÓN DE BASE DE DATOS**

Las expresiones bases de datos y bancos de datos frecuentemente se utilizan como sinónimos. Hay diversas definiciones, según se enfatizan sus características técnicas o su contenido. Para Martínez de Sousa (2004) base de datos documental, es un “conjunto de referencias bibliográficas almacenadas en soporte electrónico y estructuradas de forma que pueden ser recuperadas automáticamente.” (p. 82)

En *The Harrod's Librarians' Glossary* se define base de datos como un “conjunto de información reunida con un propósito particular para un grupo específico de usuarios. Se organiza por campos y se provee de herramientas para la manipulación de los datos” (p.190)

En el Primer seminario sobre políticas nacionales de información para la investigación y el desarrollo sobre *Bancos de Información*, organizado en México por el ex Centro de Información Científica y Humanística (CICH. UNAM) y el Servicio de Consulta a Bancos de Información (SECOBI. CONACyT) en 1987, se definió base de datos, como “... información sobre referencias [bibliográficas] almacenadas en forma electrónica con los mecanismos de búsqueda necesarios para un acceso fácil, rápido y eficiente...” y banco de datos como “...agregado estructurado de datos factuales [numéricos] que indican el dato exacto de la información que se desea conocer, almacenados en forma electrónica y con los mecanismos que permiten facilidad en la recuperación con posibilidades de interacción rápida y eficientemente” (p. 9)

Para Rodríguez Yunta una base de datos es el “conjunto de información estructurada en registros y almacenada en un soporte electrónico legible desde un ordenador. Cada registro constituye una unidad autónoma de información que puede estar a su vez estructurada en diferentes campos o tipos de datos que se recogen en dicha base de datos”. (2001, p.1)

A partir de las definiciones anteriores, y como marco de referencia en este documento se entenderá una base de datos como un conjunto de referencias bibliográficas integradas en registros, almacenadas en forma electrónica y que mediante un sistema de gestión de bases de datos, posibilita el ingreso, sistematización, almacenamiento, consulta y recuperación de la información. Cada registro está integrado por diversos campos y cada campo por diferentes tipos de datos.

El objetivo principal de una base de datos es satisfacer las necesidades de información de usuarios específicos, ofreciendo el acceso fácil, preciso y oportuno, a la información que contiene. Para lograr este objetivo se realizan acciones de acopio, selección, análisis documental, sistematización, almacenamiento, organización y difusión de la información. La tecnología actual facilita que las bases de datos incluyan el texto completo de los documentos

## **TIPOS DE BASES DE DATOS**

Las bases de datos se pueden clasificar de acuerdo a diferentes criterios, entre ellos: por el contenido temático, origen institucional, cobertura geográfica, tipo de acceso, fuentes de financiamiento, tipo de fuentes analizadas, etc.

Lancaster (1996) y Rodríguez Yunta (2001), clasifican y definen las bases de acuerdo a los criterios anteriores.

- ✓ *Por el tipo de fuentes:*
- ✓ Bibliográficas: catálogos de libros, revistas, tesis, literatura gris,
- ✓ De imágenes: fotografías, pinturas, museos
- ✓ De documentos sonoros y audiovisuales.
- ✓ *Por el nivel de procesamiento documental:* referenciales, referenciales con resumen y referenciales con resumen y texto completo
- ✓ *Por el contenido temático:* multidisciplinarias y especializadas
- ✓ *Por la cobertura geográfica:* nacionales, regionales, internacionales
- ✓ *Por el origen institucional:* públicas, privadas y de organismos sin fines de lucro
- ✓ *Por el tipo de acceso:* gratuitas o pagadas
- ✓ *Índice de citas.*

En este trabajo se abordan las bases de datos bibliográficas referenciales con resumen, y con posibilidad de ligas al texto completo, especializadas en educación, cuyas fuentes pueden ser revistas científicas y técnicas, producidas por un organismo público y por lo tanto con acceso gratuito.

## **CARACTERÍSTICAS DE LAS BASES DE DATOS**

Las bases de datos tienen algunas características comunes, entre ellas podemos mencionar las siguientes:

- ✓ *Índice de citas.*
- ✓ *Accesibilidad a múltiples usuarios.*
- ✓ *Capacidad de formar subconjuntos:* cualquier registro puede agruparse con otros que tengan características similares.
- ✓ *Flexibilidad para cambios estructurales:* en una base de datos es posible modificar su estructura, agregar campos nuevos, incrementar nuevos registros, pasar de acceso privado a público, cambiar de un software a otro.
- ✓ *Control interno para evitar redundancia:* con una metodología de trabajo documental adecuada y con un gestor de bases de datos muy bien desarrollado, se pueden generar catálogos de autoridad de temas, autores, títulos de las fuentes etc. y se puede evitar la duplicación de registros.

## **PLANEACIÓN, DISEÑO Y DESARROLLO DE BASES DE DATOS**

La metodología para desarrollar una base de datos, se puede dividir en tres fases principales: planeación, diseño y desarrollo.

En la fase de planeación se establece el modelo de la base, que implica analizar los siguientes apartados: justificación, objetivos, usuarios, alcance, cobertura temática, geográfica, temporal e idiomática; tipo y soportes de las fuentes documentales a incluir, nivel de procesamiento documental: referencial, referencial con resumen, referencial con resumen y con texto completo, productos y servicios que se proporcionarán, definición de requerimientos de insumos, equipo, software y recursos humanos

El diseño, comprende la selección del sistema informático o Sistema de Gestión de Bases de Datos, el diseño de la estructura técnica de la base que a su vez incluye la definición de

etiquetas y campos, elaboración de formatos de captura, definición de campos a indizar, formatos de visualización y despliegue, pruebas de captura y consulta de la información,

La última es la fase de operación o desarrollo de la base, implica la contratación y capacitación del personal, compra de equipo, elaboración de los criterios, políticas y estrategias para la localización, adquisición y selección de las fuentes a analizarse, el diseño de un sistema de control de calidad, elaboración de manuales técnicos, de políticas y procedimientos e implementación de un programa de promoción y difusión.

## **PLANEACION**

Al elaborar un proyecto de creación de una bases de datos es importante reflexionar sobre algunos aspectos, entre ellos:

¿Cuál es el objetivo de la base de datos?, ¿Llena una necesidad institucional?, ¿Complementa otra base ya existente?, ¿Vale la pena desarrollar una base con sólo los productos institucionales, o estos podrían ser enviados para su integración a otras bases de datos?, ¿Tiene la institución los recursos necesarios de equipo y el personal profesional para sostener su producción?, ¿Es prioritario este gasto en función de los objetivos y políticas institucionales?, ¿Es de utilidad para la comunidad?.

Con referencia a los usuarios, cabe pensar ¿Quiénes son los potenciales usuarios? ¿Pertencen a la institución, al sector privado o al gobierno?, ¿Qué tipo de información necesitan?, ¿Qué servicios requieren?, ¿Qué otras instituciones podrían interesarse en la base?, ¿Habría la posibilidad de intercambiar la información con otras bases?, ¿Se podrían generar subproductos?. ¿Las bases de datos que existen en el mercado, incluyen la información que se pretende recopilar y analizar, en qué porcentaje?, ¿La base de datos que proponemos puede ofrecer algún valor agregado?, Por ejemplo resumen de los documentos o el texto completo?

Una investigación de mercado de los productos documentales similares existentes puede validar los objetivos y metas planteadas y corroborar si es necesario o no generar el producto que se plantea, aunque con frecuencia este tipo de proyectos responden a intereses políticos e institucionales de las autoridades en turno más que a intereses académicos.

Si se ha valorado la necesidad de desarrollar la base, porque ésta llena un vacío en cuanto al contenido temático, geográfico o idiomático, pero no se cuenta con los recursos suficientes, podría plantearse la posibilidad de autofinanciamiento. En este caso habrá que presupuestar el costo del desarrollo y analizar costos de consulta, venta de servicios y subproductos.

Si la base de datos se genera para un nicho específico de usuarios, se deben considerar las características propias de este grupo. En el área educativa por ejemplo, el perfil del usuario es heterogéneo: investigadores, docentes, estudiantes, tomadores de decisiones, documentalistas, público en general, que a su vez corresponden a niveles académicos y económicos distintos.

Para realizar un estudio de usuarios se puede recurrir a diversas metodologías ya probadas, por ejemplo la de Santaella Ruiz o a técnicas de marketing, como la técnica de grupos focales. García-Morales recomienda "conocer las expectativas de nuestros usuarios y saber prestar un servicio que cubra esas expectativas es un factor decisivo para mantener a nuestros clientes y en definitiva para la pervivencia del sector" (p-12).

Es importante también analizar otras experiencias en el desarrollo de bases de datos similares, productos que generan, servicios que ofrecen, costos, software que utilizan etc..

## **DISEÑO DE LA ESTRUCTURA**

Para diseñar la estructura de la base de datos, es imprescindible haber seleccionado previamente el software a utilizar, es decir el Sistema de Gestión de Bases de Datos (SGBD), que es el conjunto coordinado de programas, que permite la creación y explotación de la base de datos, garantizando su seguridad.

Se debe tener presente que hay diferencias en los SGDB en cuanto a sus objetivos y funciones, por lo que es aconsejable elegir un software desarrollado específicamente para bases de datos bibliográficas como el MICRO CDS/ISIS, (Computerized Documentation Systems/Integrated Set for Information System), ALEPH, Inmagic, Knossys. Por ejemplo para el caso del IRESIE se seleccionó el MICRO CDS/ISIS por múltiples factores, entre otras razones porque es gratuito, producido por la UNESCO, para el tratamiento de información bibliográfica, que permite la indización con diversas técnicas y por lo tanto mayor precisión

en la recuperación de la información. Es compatible con el formato internacional ISO 2709 para transferencia de información, es utilizado para desarrollar bases de datos bibliográficas especializadas en cientos de bibliotecas de México y América Latina. Se puede instalar en todas las máquinas que se requiera, sin problemas de licencia; no se necesita de equipo muy caro, la versión windows mantiene la compatibilidad con la base de datos diseñada en el sistema operativo DOS, permite migrar versiones anteriores, realizar consultas por medio de interfáz gráfica, se puede trabajar en red y con aplicaciones para CD ROM, e incluir textos completos, tiene un sistema de indexación rápido y eficiente que facilita la recuperación selectiva.

En este trabajo se describen los lineamientos generales para el diseño de una base de datos que podría aplicarse en CDS/ISIS, software que no tiene un formato predefinido, lo que permite que cada usuario defina el suyo de acuerdo a sus necesidades. Para optimizar el proceso de diseño se sugiere tener a la mano el Manual del ISIS.

Como se dijo antes, una base de datos es un conjunto de registros, interrelacionados de manera lógica, divididos en unidades más pequeñas llamadas campos y subcampos, por ejemplo: autor, título de la revista, volumen, número, año etc.

El primer paso para diseñar la estructura de la base es el análisis detallado de los elementos básicos que la van a integrar. Para la selección de las etiquetas, los campos y subcampos se debe tomar en cuenta los siguientes aspectos:

- El tipo de documentos: libros, revistas, tesis, literatura gris, videos etc.
- El soporte físico: impresos, magnéticos, digitales etc.
- El tratamiento documental que se va a realizar, sólo referencial, referencial con resumen, referencial con resumen y texto completo.
- La localización de las fuentes: en la institución generadora o en varias instituciones.

De acuerdo al tipo de documentos, el soporte físico, nivel del tratamiento documental, necesidades detectadas en el estudio de usuarios, se elaborará la tabla de definición de etiquetas, campos y subcampos que en ISIS se denominan Tabla FDT, que contienen:

- ✓ **Elementos de descripción bibliográfica:** autor, institución de adscripción del autor, título del documento, título de la fuente, pié de imprenta, (lugar de publicación, editor, y

fecha de publicación) año, vol., no. parte, época, páginas, descripción técnica, serie, clasificación, etc.

✓ **Elementos referentes al contenido:** descriptores, identificadores, palabras clave, resumen, texto completo, número de referencias, tipo de documento (artículo, ensayo, documento histórico, entrevista etc.)

✓ **Elementos de control administrativo:** fecha de captura, nombre del documentalista, fecha de revisión, número de control, localización del documento.

Se diseñan las *hojas de trabajo* para la captura de la información, agregando ayudas para facilitar la descripción, homogenizar la captura y normalización de la información.

Se determinan los *campos a indizar*, es decir la tabla FST, en la que se definen los elementos recuperables, considerando diversos puntos de acceso a través de campos invertidos como: autor, título, fuente, pié de imprenta, descriptores, identificadores, el resumen, algunos campos de control administrativo. Se recomienda analizar y seleccionar adecuadamente las técnicas de indizado, para facilitar la recuperación con diversos fines, tanto para el control de calidad de los registros, control administrativo de producción y sobre todo pensando en los usuarios y las modalidades de consulta por índices de autor, título del documento, título de la fuente, país de procedencia, año etc., también para combinaciones en la búsqueda libre, con operadores booleanos (y, o, no), mediante estrategias de truncación, (Mexic\* ) o adyacencias para permitir búsquedas relacionadas más específicas y precisas.

Se diseñan también previamente varios formatos de visualización y despliegue para la consulta en línea, impresión e intercambio, en concordancia con las necesidades de los usuarios, los subproductos esperados y los servicios que se pretender establecer. Por ejemplo si se van a editar índices impresos temáticos, autores, títulos de las fuentes, tipo de documentos, centros en donde se localizan u otros. Por ejemplo, se diseñan formatos diversos de salida: a) **completo** que contenga todos los elementos de una referencia bibliográfica, incluyendo el resumen y una liga al texto completo; b) **corto**, que contenga los elementos mínimos de una referencia, o sólo de títulos, autores; de tipo de documentos, c) formato con etiquetas para revisión, etc.

**Llaves de seguridad.** Para evitar alteraciones de la información por parte de personal no autorizado, se deben mantener mecanismos de seguridad, mediante llaves de acceso o

passwords. No olvidar hacer respaldos de la información procesada diaria o semanalmente, de acuerdo a la periodicidad de la actualización de la base de datos. Se recomienda, también, tener una copia del diseño original de la base para cuando se requiera su instalación en nuevas computadoras.

## **NORMALIZACIÓN**

Una norma de acuerdo a la ISO (International Estándar Organization), es una especificación técnica, elaborada con la cooperación y el consenso de todos los interesados, basada en los resultados consolidados de la ciencia, la tecnología y la experiencia. Es aprobada por un cuerpo académico reconocido a nivel nacional, regional o internacional.

La normalización en el campo de la documentación juega un papel muy importante, e incide en todas y cada una de las etapas del ciclo de la información, desde su generación, edición, sistematización y la transferencia. Para el procesamiento documental existen diversos instrumentos de normalización, tales como: el formato MARC que define un sistema de códigos para identificar los datos de registros bibliográficos, los sistemas de clasificación, (Decimal Universal, Decimal Dewey, Clasificación de la Biblioteca del Congreso, etc.) tesauros, (UNESCO, Tesoro Europeo de la Educación, ERIC), Vocabularios controlados (IRESIE), la norma UNE 50-103-90 para la elaboración de resúmenes, los códigos ISSN e ISBN para identificar títulos de revistas o títulos de libros; la norma Z39.50 para recuperar y transferir registros bibliográficos en formato MARC, entre otras.

En una base de datos, los elementos que integran cada uno de los campos y registros, deben describirse de manera consistente y normalizada con el fin de que la consulta y recuperación de la información sea exitosa, y para que se pueda intercambiar la información con facilidad a nivel nacional o internacional. Por ello es imprescindible que se adopten normas para la catalogación de los documentos, clasificación, registro y transferencia de la información. (Garduño 1996).

## **DESARROLLO DE LA BASE**

Una de las primeras acciones para el desarrollo de una base de datos es elaborar por escrito las políticas y criterios para la obtención y selección de las fuentes a incluir. Por ejemplo para una base de datos de libros, definir si se incluirán todos los libros que llegan al centro de documentación o biblioteca o únicamente aquellos que tienen determinado nivel académico

y/o cobertura temática, geográfica, temporal e idiomática, quizá aquellos que fundamentan ciertos programas institucionales. Estos lineamientos estarán en función de los objetivos, metas e infraestructura disponible y sobre todo a las necesidades, tipo y nivel de usuarios al que esté dirigida la base: académicos, técnicos, o al público en general.

Se pueden definir estrategias para detectar nuevos títulos, por ejemplo a través de la revisión sistemática de reseñas de libros, de anuncios de nuevos libros o revistas, exploración de catálogos, exploración de páginas Web de instituciones, visitas a librerías físicas y virtuales, etc. Se recomienda mantener un control adecuado de los ingresos de materiales, sobre todo si se reciben por compra o canje. También deben establecerse políticas por escrito para la organización y conservación de los documentos originales que respalden a futuro la disponibilidad de la información contenida en la base, con el fin de que no se descarten dichos materiales, ni se cambien de clasificación.

## **ANÁLISIS DOCUMENTAL**

La fase esencial en la construcción permanente de una base de datos es el análisis documental. Coll-Vinent y Bernal Cruz (1990) definen el análisis documental como "la operación que consiste en extraer los elementos informativos de un documento original a fin de expresar su contenido en forma abreviada" (p.108)

Para que el producto de este proceso sea de calidad, se requiere que el personal que lo realiza tenga conocimientos especializados sobre la temática y experiencia en procesos documentales. Si no se analiza bien el documento, éste no se puede recuperar.

El análisis documental comprende dos acciones: la descripción bibliográfica y catalográfica del documento y el análisis de su contenido.

La *descripción bibliográfica* implica transcribir los elementos de identificación del documento ya sea sólo con los datos obligatorios o ampliada con elementos opcionales, aplicando las normas seleccionadas.

El *análisis del contenido* involucra a su vez :

a) *Elaboración de un resumen*, que es la representación lo más fiel posible del contenido de los documentos. Ofrece al usuario elementos de juicio adicionales para la selección del

documento o para su descarte. No todas las bases de datos incluyen resumen, aunque es lo más recomendable porque se le ofrece al usuario un elemento importante para que tome decisiones más adecuadas al seleccionar los documentos.

Existen varios tipos de resúmenes, los más utilizados por los documentalistas del IRESIE son: el *indicativo* que describe los temas principales y secundarios abordados en un documento, sin incluir crítica alguna; el *informativo o analítico* que reseña de una forma más profunda los aspectos significativos y relevantes de un artículo. Generalmente se utiliza para los reportes de investigación, incluye el o los objetivos del estudio, el método de investigación aplicado, la muestra, principales procedimientos, los resultados, y algunas conclusiones del estudio (IRESIE, Manual de políticas y procedimientos)

Para Pinto Molina el resumir es “un proceso de recreación a escala reducida del documento original, una difícil y compleja operación creativa, o mejor dicho re-creativa, cuyo carácter se revela ante cualquier intento normalizador”, (1991, p.159) exige del documentalista conocimientos y habilidades para leer, mirar, comprender, analizar, sistematizar y sintetizar en función de las necesidades, tipo y nivel de usuarios y desde luego la naturaleza y soporte de los documentos.

La calidad del resumen por lo tanto, depende de la capacidad del documentalista para reflejar el contenido de un documento en pocas palabras, y también de la estructura misma del documento.

Ante la creciente cantidad de información y las características de los documentos en educación, es saludable pensar en la necesidad de que los autores, sean capacitados para la elaboración normalizada de resúmenes, de tal manera que los incluyan al inicio de cualquier publicación ya sea libro, artículo de revista, ponencia u otro texto, facilitando la labor del documentalista, asegurando la pertinencia del resumen y optimizando el acceso al documento. Los editores pueden jugar un papel central para establecer normas rigurosas para la elaboración de los distintos tipos de resúmenes de acuerdo a la naturaleza y contenido del documento, ya sea un artículo teórico, de revisión, un reporte de investigación, un manual, una monografía etc. (García y Torres, 2006)

## INDIZACIÓN

La UNESCO define indización como un proceso que consiste en describir y caracterizar un documento con la ayuda de representaciones de los conceptos contenidos en dicho documento. Para Neet indizar es “analizar los documentos y aislar en la riqueza del lenguaje natural empleado por los autores, todos los conceptos esenciales o pertinentes que deben ser retenidos en previsión de búsquedas posteriores” (cit por Pinto Molina, p. 127).

Una definición más general de indización y que se puede aplicar a las bases de datos, es la anotada por Atherton, “proceso de identificación y asignación de rubros, descriptores o encabezamientos de materia relativos a un documento, de modo que su contenido sea conocido y el índice creado pueda ayudar a recuperar elementos de información” (1978, p.184)

Con base en la anterior definición, me parece que el término indización debe reemplazarse por *Asignación de descriptores, identificadores o palabras clave*, proceso que consiste en conceptualizar el contenido del documento y traducirlo a un lenguaje normalizado como los descriptores e identificadores, o a un lenguaje libre expresado en palabras claves, con la finalidad de crear índices, los cuales son generados en la actualidad por el sistema de gestión de bases de datos.

Este proceso supone un análisis del contenido del documento, basado en el conocimiento de la disciplina, la estructura de la base, las necesidades del usuario y el vocabulario normalizado que se utiliza. De la fidelidad, consistencia y especificidad en la asignación de términos, depende la pertinencia y precisión en la recuperación de la información, ya que el lenguaje del análisis es el puente entre la información que contiene el documento y el usuario que consulta la base.

Siendo este proceso subjetivo, la posibilidad de reflejar exactamente el contenido de un documento no es factible. Se han llevado a cabo ejercicios en los que la coincidencia en la asignación de términos a un mismo documento, entre varios documentalistas llega al 40 o 50%, dependiendo de la disciplina. Saracevic, T. (1993) afirma, que aún entre expertos en el campo de estudio, la concordancia en la terminología de la indización de un mismo documento es del 40% en promedio.

Esto refleja el difícil papel del documentalista en la complicada tarea de interlocutor entre el contenido y la naturaleza del documento y la necesidad del usuario, sobre todo con este cambiante escenario de las tecnologías que hace mucho más complicado el proceso del análisis documental y la importancia de recuperar de manera puntual. (Torres y García, 1999)

Hay diversos sistemas de asignación de términos, entre ellos están:

Por **descriptores**, se basa en conceptos, es más preciso y elimina el riesgo de dispersión, en este caso se deberá seleccionar el tesoro apropiado de acuerdo a la temática de la base. Por ejemplo para el área educativa existen el Tesoro de Educación de la UNESCO, el Tesoro Europeo de la Educación, el del ERIC. Una desventaja del uso de tesauros es que no se actualizan con la velocidad en que van apareciendo nuevos términos y conceptos, y que además, en general, los usuarios que consultan las bases utilizan lenguaje natural.

Por **unitérminos**, basados en las palabras significativas del título, es decir en lenguaje natural. Una desventaja es que no siempre el título representa en su totalidad el contenido del documento y en ocasiones éste ni siquiera corresponde al contenido.

En la base de datos IRESIE, para la asignación de descriptores se utiliza un Vocabulario Controlado, cuyos vocablos, en gran parte, fueron tomados del Tesoro de la Educación de la UNESCO y otros traducidos del ERIC; sin embargo, frecuentemente se ingresan nuevos términos extraídos del contenido de los artículos.

Sugiero que los términos asignados a los documentos amplíen su alcance con el fin de reflejar, además de las categorías temáticas del contenido, otras características como: el nivel de profundidad, contexto disciplinario, orientación teórica, tipo de abordaje, teniendo siempre presentes las necesidades de los usuarios. (García y Torres, 2001) .

## **CONTROL DE CALIDAD DEL ANÁLISIS DOCUMENTAL**

Este aspecto es de trascendencia para mantener la calidad de la base. Consiste en la revisión sistemática de los diversos elementos y acciones que se llevan a cabo en el proceso del análisis documental, con base en las políticas de selección y aplicación de las normas. Se cotejan los datos de la descripción bibliográfica: autor(es), título del documento, título de la fuente, país, volumen, número, año de edición etc., para detectar, omisión de datos en campos indispensables, inclusión errónea de información, errores ortográficos, equívocos

gramaticales de sintaxis, como frases mal construídas, incorrecciones semánticas, imprecisión en la asignación de descriptores, identificadores, claridad de los resúmenes. Se revisan los diferentes índices de autores, títulos del documento, título de la fuente, descriptores, identificadores, etc.

Es importante tener presente que para la consulta de la base hay equivocaciones de más trascendencia que pueden distorsionar la recuperación de la información, por ejemplo si está mal escrito el apellido de un autor, el título del documento, o la errónea asignación de un descriptor, así mismo por razones de alfabetización en la indización, repercute más la errata en la raíz de la palabra, que aquella que está al final.

En la base de datos IRESIE hemos encontrado que el intercambiar, sin nombre, los registros con sus resúmenes y descriptores entre los diversos documentalistas, permite detectar errores de redacción y falta de claridad tanto en los resúmenes como en la asignación de los descriptores, que pueden ser pensados desde la formación personal con una connotación diferente. El contrastar las diversas conceptualizaciones facilita ir conformando una base homogénea en la comprensión del vocabulario empleado en la base.

### **RECURSOS HUMANOS:**

El equipo humano mínimo que requiere una base de datos está constituido por: documentalistas, bibliotecarios, informáticos y personal de apoyo administrativo. Es elemental cuidar el perfil profesional de cada uno de los integrantes del equipo de trabajo, de acuerdo a las funciones que desempeñen en la producción de la base. Se requiere que todos los integrantes del equipo, además del manejo de contenidos y capacidades técnicas respectivas, conozcan y compartan la filosofía del proyecto, y desde luego estén dispuestos a mantenerse actualizados.

El bibliotecario responsable de la adquisición de los fondos documentales, se caracterizará por conocer el mercado editorial correspondiente, sea de libros, revistas, videos, películas, fotografías, etc. ser organizado, hábil para las relaciones interpersonales e interinstitucionales. Entre sus funciones está el adquirir mediante compra, canje o donación los materiales documentales correspondientes así como mantener catálogos o directorios de los distribuidores que ofrecen mejores servicios y precios, tener un control del ingreso de los documentos a la base.

Dependiendo del tamaño, características y ubicación del fondo documental, este mismo personal será responsable de la organización y conservación del acervo documental, soporte físico de la base de datos.

Los documentalistas, que llevan a cabo el proceso de análisis, asignación de descriptores, identificadores o palabras clave y elaboración de resúmenes (si es el caso), además de ser especializados en el contenido temático de la base, conviene que se capaciten en el manejo del software y en el conocimiento de las normas internacionales para el proceso documental, tales como las Reglas Angloamericanas de Catalogación, Formato Marc, uso de tesauros, desarrollo de vocabularios controlados. Es deseable que tengan cualidades como: buena memoria, ser organizados, habilidades intelectuales para la lectura, análisis, síntesis y redacción y desde luego facilidad para interrelacionarse y trabajar en equipo.

El revisor del control de calidad del proceso documental (si lo hubiere) conviene que tenga un perfil profesional relacionado con la temática, que desarrolle lo que comunmente se denomina “ojo clínico”, posea cualidades intelectuales para examinar sistemática y minuciosamente cada uno de los datos y elementos de las referencias bibliográficas, descripciones técnicas y mejorar si es el caso la redacción de los resúmenes. En el caso del IRESIE, por ejemplo, que actualmente sólo cuenta con dos documentalistas y un coordinador, se hace la revisión periódica sólo de una muestra de registros.

El informático, responsable del mantenimiento técnico de la base, además de conocer el software, conviene que maneje programas de paquetería complementaria y resuelva los problemas técnicos que se vayan presentando. En la mayor parte de las instituciones el apoyo de personal de cómputo es ocasional y no es adscrito al staff.

Finalmente el personal que ofrece los servicios como: búsquedas especializadas, recuperación de documentos o disseminación selectiva de información, venta y difusión de productos, es necesario que comulgue con una filosofía de trabajo basada en la satisfacción de los usuarios, sea paciente, tenga disposición de servicio, habilidad para comunicarse y conozca a profundidad los antecedentes, objetivos, políticas, cobertura, alcance y contenido de la base de datos propia y otras relacionadas, para que puede canalizar al usuario a otras fuentes de información, si así se requiere, después de hacer un diagnóstico de sus

necesidades. Es deseable que los mismos documentalistas ofrezcan algunos de estos servicios a los usuarios con el fin de que haya retroalimentación.

El número y perfil profesional del personal para desarrollar la base, depende del nivel de análisis, la cantidad de documentos, el tipo de las fuentes y la cobertura temática e idiomática de la misma. Por ejemplo, si la base de datos incluye diversos tipos de documentos y soportes documentales, está en diferentes idiomas, su nivel de análisis es completo con resumen, los productos esperados son especializados, conviene que el personal sea más numeroso y su perfil profesional más especializado, que si el nivel del análisis es únicamente referencial.

La generación de una base de datos exige trabajo en equipo, cada profesional se responsabiliza de distintas fases de un proceso sistemático para alcanzar los objetivos y metas comunes, por ello se recomienda una metodología participativa de todos los miembros del equipo, mediante la organización de sesiones periódicas de trabajo conjunto, que propicien la comunicación entre el personal técnico, directivo y de apoyo administrativo, a partir de un ambiente laboral que facilite la cooperación, se unifiquen criterios en la selección de los documentos, asignación de descriptores, se conozcan las temáticas y fuentes documentales más demandadas por los usuarios, se resuelvan los problemas que se vayan presentado y se propicie además el crecimiento profesional y la toma de decisiones.

Durante el desarrollo del IRESIE, principalmente en las etapas en las cuales se contó con apoyos financieros externos y el equipo de documentalistas estaba integrado por cuatro o cinco personas se experimentó con éxito esta forma de trabajo.

## **SELECCIÓN DEL HARDWARE Y SOFTWARE**

Las características del hardware y el software se definen en concordancia con el sistema de bases de datos a desarrollar, así como con los recursos económicos disponibles. Si el equipo de cómputo destinado para la producción de la base de datos es de gran capacidad, se tendrá ganado un elemento que contribuye a la calidad. No es lo mismo una computadora 486 que una pentium IV, ni es lo mismo tener una sola computadora que cinco o diez trabajando en red para la producción de una base de datos mediana. Generalmente a nivel institucional para la selección del hardware se cuenta con personal especializado que colabora en este rubro.

Es conveniente evaluar los SGBD con el fin de escoger el más adecuado en concordancia con el equipo disponible, el tipo y alcance de la base a diseñar, la compatibilidad con otras bases, el presupuesto que se tenga, la facilidad de uso, el apoyo técnico local, idioma en el que esté desarrollado. Para minimizar errores en la selección se aconseja recabar información de otras instituciones sobre sus experiencias acerca de la paquetería.

Para el usuario final es deseable que el software facilite la eficiente exploración y explotación del contenido de la base, para realizar búsquedas libres básicas, búsquedas avanzadas precisas mediante operadores booleanos de intersección, conjunción o negación de términos, de adyacencia, aproximación, que se puedan combinar conjuntos, guardar y repetir estrategias de búsqueda, tener interfaces gráficas que permitan visualizar los índices y los diferentes campos de la base, con un lenguaje de recuperación fácil, con ayudas y mensajes de error.

Para control administrativo es deseable contar con programas para detectar registros duplicados, inconsistencias al asentar datos de los diferentes campos, hacer correcciones ortográficas.

## **ACCESO Y DIFUSIÓN**

La base de datos debe estar accesible para su consulta a través de diversos medios, en CD-ROM, en Internet a través del portal de su institución o un sitio propio y quizá producir algunos productos impresos.

La difusión de la base, sus productos y servicios, se puede orientar hacia un nicho de usuarios específico o hacia un público más amplio, a través de medios diversos como: publicación de artículos en revistas especializadas, conferencias o ponencias en eventos académicos, demostraciones a grupos de usuarios, anuncios en medios impresos o electrónicos, directorios de bases de datos, gacetas, carteles, trípticos, etc.

La comercialización de los productos, si no es posible hacerlo institucionalmente, se dejará en manos de una empresa distribuidora de bases de datos. Este rubro dependerá de las políticas institucionales y del tamaño y características de la base.

## **RETROALIMENTACIÓN POR PARTE DE LOS USUARIOS**

Para saber si se están cumpliendo las expectativas de los usuarios o conocer las inconsistencias, fallas en el análisis documental, errores detectados, fortalezas o debilidades de los productos generados, limitantes de los formatos de visualización o impresión, limitantes de los servicios, necesidades de nuevos servicios etc. es conveniente mantener un sistema de retroalimentación, mediante instrumentos de evaluación como estadísticas de servicios, encuestas, cuestionarios, entrevistas, buzones de sugerencias, entre otros, ya sea aplicando alguno de estos instrumentos en forma directa a usuarios que consultan personalmente la base o a través de la red.

## **HERRAMIENTAS NORMATIVAS, TÉCNICAS Y BIBLIOGRÁFICAS DE APOYO**

Es necesario mantener en el área de trabajo de los documentalistas obras de consulta como: diccionarios de la lengua, de sinónimos y especializados, enciclopedias especializadas, léxicos, atlas, tesauros, manuales técnicos del software, hardware, Reglas de Catalogación Angloamericanas, el Formato Marc, algunas normas ISO para la elaboración de resúmenes y otras herramientas de trabajo intelectual que deben mantenerse siempre a la mano para solucionar las dudas que se presenten.

El **Manual de Políticas y Procedimientos**, es otro de los instrumentos básicos para el trabajo del equipo. Contiene las políticas generales y las normas específicas, los métodos, procedimientos y estrategias a seguir en cada una de las tareas.

Su contenido abordará, al menos, los rubros siguientes:

- Antecedentes, objetivos, metas, usuarios, productos y servicios de la base de datos
- Estructura y características técnicas de la base, incluye la tabla de definición de los campos y etiquetas, formatos de captura u hojas de trabajo, formatos de visualización e impresión.
- Políticas de selección de los materiales y formas de adquisición.
- Procedimientos para la descripción bibliográfica, el análisis de contenido, elaboración de los diferentes tipos de resúmenes, codificación y captura.
- Procedimientos para el control de calidad, cotejo y validación de los datos
- Procedimientos para elaborar los productos
- Política y procedimientos de los servicios.
- Políticas de difusión, comercialización y costos de productos y servicios.

## CONCLUSIÓN

Para concluir quisiera hacer un símil con un equipo de futbol, en el que un jugador mete el gol, pero éste es el resultado de un trabajo conjunto del equipo. De igual forma, la base de datos es el producto final pero ésta es el resultado de una serie de acciones llevadas a cabo de manera coordinada para detectar, adquirir y seleccionar las fuentes documentales, realizar el tratamiento documental de análisis y sistematización de la información en la computadora, organizar y conservar los documentos originales para su posterior localización y ofrecer servicios y productos para satisfacer necesidades de los usuarios.

En este documento he tratado de plasmar las experiencias vividas durante el largo período de 28 años de desarrollo del IRESIE que con muchos altibajos, al pasar por cuatro dependencias de adscripción y estar con frecuencia al borde de desaparecer, ha salido a flote y logrado mantenerse por muchos años, no sólo la base de datos, sino fundamentalmente el acervo físico que le respalda, a pesar de estar éste disgregado en nueve bibliotecas.

## BIBLIOGRAFÍA

- Atherton, P. (1978). *Manual para sistemas y servicios de información*. París: UNESCO.
- Coll-Vinent, R. y Bernal Cruz, F. (1990). *Curso de documentación*. Madrid: Dossat.
- García Colorado, C. y Torres Verdugo, M.A. (2006). *Complejidad y organización de la información en el área educativa: propuesta de documentos orientadores como auxiliares del usuario de la información*. Mimeografiado 14 p.
- García-Morales Huidrobo, E. (1995). Aportaciones de la gestión de calidad a bibliotecas y servicios de documentación. *Revista Española de Documentación Científica*, 18 (1), 9-18.
- Garduño Vera, R. (1996). *Modelo bibliográfico basado en formatos de intercambio y en normas internacionales orientado al control bibliográfico universal*. México: UNAM. Centro Universitario de Investigaciones Bibliotecológicas, (Serie monografías, 9).
- Lancaster, F. (1996). *Lineamientos para la recopilación de bases de datos*. México: UNAM. Centro Universitario de Investigaciones Bibliotecológicas.

*Manual de políticas y procedimientos del IRESIE* (1998). México: UNAM. CESU (documento técnico, mimeografiado)

Martínez de Sousa, J. (2004). *Diccionario de bibliología y ciencias afines*. 3ª. ed. muy aumentada. Gijón, Asturias: Trea.

Pinto Molina, M. (1991). *Análisis documental: fundamentos y procedimientos*. Madrid: Eudema Universidad.

Primer Seminario sobre Políticas Nacionales de Información para la Investigación y el Desarrollo (1987). *Bancos de Información*. México: Centro de Información Científica y Humanística (CICH. UNAM), Servicio de Consulta a Bancos de Información (SECOBI. CONACyT) (Documento de trabajo, No. 3, 21 p.).

Prytherch, R. (comp.) (1995). *Harrod's Librarians' Glossary*. 8th. ed. England: Gower.

Rodríguez Yunta, L. (2001). Bases de datos documentales: estructura y principios de uso en Maldonado, Angeles (Coord.). (2001). *La Información especializada en internet*: Madrid: CINDOC, [<http://www.unav.es/dpp/documentacion/proteger/lryunta.pdf>] consultado 29/11/06

Ruiz Velasco Romo, E. (1996). Elaboración de una base de datos analítica: una experiencia en micro CDS/ISIS. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, (Manuales; 19)

Santaella Ruiz, R. D.(2005 ). Metodología de estudios de usuarios de la información: estudio de casos en la Administración Pública. *Revista TEXTOS de la CiberSociedad*, 5 [<http://www.cibersociedad.net/textos/articulo.php?art=61>] consultado 26/11/06).

Saracevic, T. (1993). Issues of Quality in information services by intermediaries: application of total quality management (TQM). *Conferencia presentada en el Segundo Foro Latinoamericano de Información LATINBASE 93*. México: Fil Guadalajara.

Torres Verdugo, M. A. (2000). Elementos de calidad en el desarrollo de bases de datos de documentos audiovisuales. (pp.145-175). *Bancos de imágenes audiovisuales con fines educativos*: México: Dirección General de Televisión Educativa. Centro de Entrenamiento de Televisión Educativa.

Torres Verdugo, M. A. y García Colorado, C. (1999). Diseño y desarrollo de bases de datos bibliográficas: algunos factores que inciden en su calidad. *Información: producción, comunicación y servicios*, (39), 3-18.