



# BASES DE DATOS DOCUMENTALES

⇒ MIGUEL A. CORCOBADO, BELÉN GÓMEZ \*

El objetivo de este artículo es la aproximación a los Sistemas de Gestión de Bases de Datos Documentales (SGBDD), y la valoración de su uso para la informatización de fondos de bibliotecas, ya que todos ellos permiten configurar un diseño de base de datos en relación con las necesidades de cada biblioteca (base de datos de usuarios, fichas, préstamos, etcétera). No se trata de programas específicos para la gestión de bibliotecas, pero son capaces de resolver los problemas derivados del almacenamiento y localización rápida de datos en grandes volúmenes de información textual. Para empezar daremos algunas pinceladas de lo que son los SGBDD y sus ventajas, y luego veremos las características de algunos de los programas más extendidos del mercado.

## RELACIONALES O DOCUMENTALES

Las Bases de Datos Documentales son diferentes al resto de gestores de bases de datos, ya que son adecuados para información no estructurada (textual, multimedia), es decir, no dividida en campos. Además estos programas disponen de todas las funciones necesarias para el almacenamiento, indexación, consulta y modificación de información textual, necesarios en todo tratamiento documental. Sobre todo son importantes porque gestionan gran cantidad de información con un alto grado de homogeneidad, que permite que sea compartida por varios usuarios. Las diferencias más claras con los gestores de bases relacionales son:

- Los registros son documentos y no fichas o tablas.
- Los campos o atributos tienen

una longitud variable y no fija.

- Un registro puede no tener estructura al estar formado por un sólo campo.
- Permiten una recuperación libre de la información, basta con indicar el contenido sin necesidad de especificar el campo.
- Mayor velocidad de recuperación.
- Cada palabra incorporada es indexada, se forman índices en cada campo.

En resumen, puede afirmarse que para la adopción de un Sistema de Gestión de Bases de Datos para la informatización de una biblioteca o



*En una biblioteca  
o centro de  
documentación,  
es mucho más  
coherente decidirse  
por un SGBD  
documental que un  
SGBD relacional*

centro de documentación -una vez descartada la adquisición de un programa específico de gestión bibliotecaria- es mucho más coherente decidirse por un SGBD documental que un SGBD relacional.

## CRITERIOS DE SELECCIÓN

A la hora de elegir un SGBDD hay que pararse a determinar una serie de criterios, como son: criterios de almacenamiento (dependerá del volumen de datos), si proporciona un buen lenguaje de interrogación, que lleve un correcto mantenimiento de los datos (evita inconsistencias en la base), criterios de acceso y estructuración (modificación progresiva de la base, definición de ésta, forma de carga de datos, control de accesos para evitar que un acceso masivo devalúe el uso) y aspectos relacionados con la seguridad (autorizaciones a usuarios, tipos de acceso, sistemas de recuperación contra fallos).

La cualidad de rapidez de búsqueda de los SGBDD viene determinada a que el programa, cuando realiza la búsqueda, lee los textos desechando una serie de "palabras vacías" (artículos y preposiciones como mínimo), que son previamente definidas por el usuario y almacenadas en un diccionario, con lo cual almacenan una sola vez toda nueva palabra que se encuentra en un registro índice. De esta forma sólo necesita buscar la palabra en el índice y no consultar registro por registro. Sin embargo en búsquedas más complejas (dos palabras o una frase), es donde se determina la calidad del producto en relación con el tiempo de búsqueda.

En los SGBDD la terminología y los objetivos de las consultas varían entre: *referencias simples*



(búsqueda de una sola palabra o descriptor), *referencias genéricas* (búsqueda de descriptores con los comodines \*, ?), *referencias adyacentes* (descriptores multitérmino, una frase o dos palabras seguidas) y *referencias cualificadas* (los descriptores no se buscan en toda la base sino en un campo determinado).

No hay que olvidar que al decidirse por usar estos SGBDD hay que realizar un estudio previo de viabilidad del mismo, para lo que se estiman los siguientes pasos:

- Estudio de necesidades y requerimientos del sistema.
- Selección de una base de datos de acuerdo a esas necesidades.
- Una vez adquirida, estar al tanto de la evolución del producto y también del mercado, para aprovechar nuevas ventajas del mismo.

## PROGRAMAS

●  
CDS/ISIS



Programa creado por la UNESCO en 1985, con el fin de promover la normalización y el intercambio de información bibliográfica entre organismos. Cuenta actualmente con más de 5000 usuarios registrados. (En un próximo número de EDUCACION Y BIBLIOTECA, publicaremos más información sobre este programa).

**Estructura.** Conjunto de seis programas independientes accesibles desde un único menú. A través de un módulo adicional se pueden añadir programas en lenguaje Pascal con objeto de mejorar las funciones del programa o adecuarlo a las necesidades del centro. Los menús se ofrecen en versiones normalizadas en inglés, francés y español.

**Funciones.**

- Definición de la base de datos.

La estructura de la base de datos se limita a 200 campos y hay un límite global de 8.000 caracteres por campo y registro. Cada campo puede estar formado por subcampos y todos los campos son repetibles. Utiliza cuatro elementos para definir la base de datos: una tabla de definición de campo (que relaciona los campos); una tabla de selección de campos (que determina los campos que se añaden al índice y la técnica de indización a usar); los formularios de introducción de datos (diseños de pantalla para introducir datos) y el formato de visualización.

\* **Introducción de datos.** Se utiliza el formulario establecido por el usuario durante la definición de la base de datos. Se pueden establecer más de una para cada base de datos. Puede importar datos en formato MARC.

\* **Búsqueda.** Utiliza una amplia gama de búsquedas basadas en índices invertidos. Para realizar una búsqueda se puede acceder a los registros o utilizar el índice. Este se puede visualizar para examinar las entradas. Se pueden realizar búsquedas de diferentes maneras como: lógica booleana, truncamiento para buscar por raíz, búsqueda por proximidad para palabras cercanas y calificativos de campo para limitar la búsqueda a un campo determinado. Permite la elaboración de un catálogo de sinónimos.

\* **Salida de datos.** Los registros se pueden ver en pantalla, imprimir en papel o almacenar en un archivo o disco magnético.

**Hardware.** IBM PCs y compatibles con 640 K RAM y disco duro. Existen versiones para ordenadores mayores.

**Proveedor.** En España el CINDOC lo distribuye gratuitamente para usuarios sin ánimo de lucro; únicamente cobra la traducción del manual de usuario al español (6.200 pts.).

UNESCO, 7, Place de Fontenoy, 75700 París.

CINDOC, Joaquín Costa, 22, 28002 Madrid. ☎ (91) 563 54 82.



Base de datos documental, que es utilizada en muchas bibliotecas del mundo para la gestión de información. Dispone de diversos programas complementarios como el Search-Magic (búsquedas por menús para el usuario) y el MARC Adaptor, para captura de registros en formato MARC. Unas 20 instalaciones en España y más de 16.000 en el extranjero.

**Estructura.** Es posible definir tantas bases de datos como deseen. Cuenta con una serie de utilidades de ayuda en cada función, lo que hace muy sencillo su manejo.

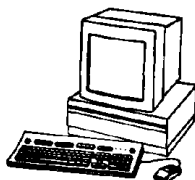
**Funciones.**

\* **Definición de la base de datos.** Ofrece campos de longitud variable y repetibles, con un máximo de 75 campos de los cuales los 50 primeros se pueden indizar. No hay límite de longitud por campo. La estructura de la base de datos se define especificando los campos y etiquetas de sus nombres, del tipo o de la indización requerida por campo, y del tipo de datos y orden de clasificación requerida. Puede manejar datos, campos numéricos, textuales y puede proporcionar ciertos esquemas de clasificación complejos como la CDU.

\* **Entrada de datos.** Los datos se pueden introducir directamente en la base de datos o alternativamente; los archivos se pueden importar en diferido. Al grabar un registro nuevo o modificado, los índices se actualizan inmediatamente. Posibilidad de integrar imágenes en formato TIFF.

\* **Búsqueda.** La búsqueda en el programa principal se puede realizar por menús o por instrucciones. Hay posibilidad de truncamiento, de búsqueda por operadores booleanos, aritméticos y por rangos. El Search-Magic permite a los usuarios inexpertos realizar búsquedas dirigidas por menús.

\* **Salida de datos.** Los registros



pueden visualizarse en los formatos que defina el usuario. Además ofrece posibilidades de creación de informes y los datos pueden clasificarse en cinco campos diferentes simultáneamente.

**Hardware.** IBM PC y compatibles, sistema operativo MS-DOS. Funcionamiento en redes Novell o Netbios compatible. 500 Kb de RAM. 2 Mb de disco duro (para el programa).

**Proveedor.** DOC 6. Tuset, 21, 6<sup>è</sup>. 08006 Barcelona. ☎ (93) 414 06 79. Comandante Zorita, 8, 1<sup>º</sup>2<sup>º</sup>. 28020 Madrid. ☎ (91) 553 52 07. Existe un Club de Usuarios de Inmagic.

**Precio.** Módulo básico 190.000 pts.

●  
KNOSYS

KNOSYS

Es una combinación de tratamiento de textos, base de datos textual y generador de informes, que integra un completo editor, registros y campos de longitud variable, y un amplio repertorio de funciones de recuperación. Acceso a los menús a través de ventanas. Existe la versión CD-Knosys, para bases de datos en soporte CD-ROM.

**Funciones.**

\* Definición de la base de datos. Los campos pueden contener entre 0 y 32.000 caracteres, lo que hace que se ocupe en la base de datos sólo el espacio utilizado por la información. Los índices tampoco están limitados en su longitud lo que significa que Knosys indexará por cada una de las palabras o fechas, números, etcétera, contenidos en los campos.

\* Introducción de datos. Identifica y trata automáticamente la información introducida, ya sea texto, fechas o números, o incluso todos ellos mezclados dentro de un mismo campo. Permite importar datos de cualquier programa de tratamiento de textos y exportarlos en formato ASCII.

\* Búsquedas. Para realizar búsquedas utiliza comandos booleanos

en castellano (O, Y, NO, EN, etcétera). También realiza búsquedas por sinónimos o términos relacionados. Contiene un diccionario de palabras vacías modificable por el usuario.

\* Recuperación de la información. Puede recuperar información no sólo por términos exactos sino también por términos relacionados.

A la hora de imprimir la información la ordena por una o varias palabras clave. Ofrece la posibilidad de extraer información al mismo tiempo que se está elaborando un informe, se está imprimiendo o realizando una búsqueda.

**Hardware.** IBM PC o compatible con sistema operativo MS-DOS 3.0 o superior y 475 Kb de RAM. Es necesaria una unidad de disco y disco duro.

**Proveedor.** Micronet. Sta. Engracia, 6. 28010 Madrid. ☎ (91) 410 50 01.

●  
BRS/SEARCH



Es un programa desarrollado para búsquedas rápidas en bases de datos, cualquiera que sea su tamaño y estructura. Incorpora un sistema de correo electrónico y conmutación de mensajes on-line, y una opción de compresión de las bases de datos.

**Funciones**

\* Definición de la base de datos. Se diseñan, se crean y se mantienen con BRS/Maint, sistema basado en menús que permite a los usuarios definir todos los parámetros. La base de datos podrá contener alrededor de unos 16.777.000 documentos, lo que supone que puede contener trillones de caracteres, el límite lo marca el hardware y el sistema operativo. En la versión micro/mini las bases de datos están compuestas por campos de longitud variable. Estos pueden ser diseñados para acomodarse a una longitud o contenido específico. Los campos de longitud variable podrán contener

hasta 65.000 caracteres.

\* Introducción de documentos. Se pueden realizar on-line o mediante la selección de la opción equivalente en el menú de BRS. Alternativamente, el usuario puede dejar un documento como fichero de entrada para realizar el proceso de carga en modo batch. Cuando se introducen los documentos existe un procedimiento de verificación automático que asegura el contenido de cada campo, según el criterio utilizado en la creación de la base de datos.

\* Búsqueda. Utiliza operadores booleanos y posicionales. En los campos numéricos pueden utilizarse operadores relacionales estándar conocidos como parámetros evaluados. Todos los operadores y los términos de búsqueda se pueden combinar de diferentes maneras. Son posibles las búsquedas con truncamientos mediante la característica Pattern-Matching que permite la recuperación de términos o palabras que comiencen, terminen o contengan una cadena de caracteres comunes.

\* Recuperación. Existen dos formas de visualizar el contenido, o una parte del mismo, de los documentos. La primera vez mediante el uso de una etiqueta especial de párrafo, que visualiza los párrafos que contengan los términos de búsqueda. La segunda mediante el uso del comando que visualiza cada una de las frases que contenga los términos de búsqueda. Es posible diseñar formatos especiales de salida, de acuerdo a las necesidades de cada usuario. Se pueden marcar los campos como "no visualizable", lo que permite realizar búsquedas, pero restringe la impresión de este campo.

**Hardware.** Cualquier terminal ASCII u ordenador personal con un software de comunicaciones. Funciona con la mayoría de los sistemas operativos. Existen versiones para grandes ordenadores y para micros bajo MS-DOS.

**Proveedor.** Servicios de Teledocumentación-Baratz. Luchana, 4-2<sup>º</sup> A-B. 28010 Madrid. ☎ (91) 593 96 12

**PUBLICIDAD**



## BASIS PLUS

# BASIS PLUS

Está diseñado para soportar grandes volúmenes de información, permitiendo su actualización en modo on-line y en diferido. Sistema que integra texto, imágenes y recuperación de objetos (gráficos, hojas de cálculo, etcétera).

**Funciones**

- \* Definición de la base de datos. BasisPlus posee una estructura modular, girando alrededor del módulo central (DM) del sistema. Este módulo es el motor del procesador de transacciones que administra los cambios en la base de datos. Aparte tiene diferentes módulos que permiten realizar varias gestiones. No hay límite en el tamaño del documento o en el formato de la información y se adapta a las necesidades del usuario.

- \* Introducción de datos. Permite gestionar documentos con diversas estructuras; desde las convencionales (datos alfanuméricos), continuas (grandes textos) y seccionadas (estructura de capítulos en formato libro) hasta los documentos que contengan formato final (negritas, distintos tipos de letra) e imágenes o gráficos.

- \* Búsqueda. Gran velocidad en las búsquedas. No necesita ningún operador, basta con el uso de una palabra, frase o concepto. Posee un módulo para entornos Windows denominado Express/W.

- \* Salida de datos. La recuperación de información es fácil y rápida, ya que está basada en el contenido de la información. Al igual que ocurre en la búsqueda, el módulo Express/W permite la salida de datos.

**Hardware.** Está disponible en una amplia gama de plataformas hardware (IBM, DEC, Fujitsu) y múltiples sistemas operativos (VMS, UNIX, DECNET).

**Proveedor.** Centrisa. Córcega, 373. 08037 Barcelona.  
☎ (93) 207 55 11.

## TEXTO WINDOWS

# TEXTO

Integrado en el sistema Windows, Texto permite realizar operaciones propias de este sistema (ventanas, cortar, pegar, etcétera) contando además con una continua ayuda en línea. Es posible a través de esta versión recuperar ficheros de antiguas versiones de Texto. Existen dos versiones más aparte de ésta, Texto (con posibilidades de aplicaciones videotex y CD-ROM) y Texto-GED (Gestión Electrónica de Bases de Datos) que permite el acceso multimedia al documento original. Actualmente existen más de 5.000 instalaciones en el mundo.

**Funciones.**

- \* Diseño de la base de datos. Permite el diseño de bases de datos de manera simple. Se pueden ir mejorando o modificando progresivamente, por lo que no es necesario hacer un diseño muy exhaustivo previo. Crea tantas bases de datos, ficheros y fichas como le permita el disco duro.

- \* Introducción de datos. Es posible modificar en todo momento los datos introducidos (ficheros, campos), pudiendo crear de 10 a 20 índices por fichero, con inversión de uno o varios campos en el mismo índice. Se pueden realizar transferencias de ficheros entre ordenadores de distinto tipo. También tiene interrelación con programas de tratamiento de textos.

- \* Búsqueda y recuperación. Recupera la información mediante interrogación multicriterio sobre uno o varios campos y uso de operadores booleanos, truncados, búsquedas por etapas históricas, etcétera. Como particularidad permite asociar en una misma orden una función de selección con una de corrección, edición, creación o borrado. Es posible consultar por tesauros, tablas y ficheros de autoridades, además de poder seleccionar los términos a consultar mediante los índices. La visualización de las fichas por ac-

ceso directo es por el número de orden, por lotes o por consultas; crea listados ordenados según cualquier criterio, permite editar el formato de las fichas.

**Hardware.** Existen versiones bajo DOS, con PC compatible; versión para red local Novell, versión UNIX y versión Windows 3.0, con procesador 386, disco duro y 4 Mb de RAM.

**Proveedor.** Chemdata Ibérica. Alcalá, 101, 2º. 28009 Madrid. ☎ (91) 435 22 02/63. Proporciona ayuda telefónica, manuales, además de varios cursos formativos a distintos niveles.

## WINDOWS PERSONAL LIBRARIAN

# WPL

Programa preparado para gestionar grandes volúmenes de información textual y que incluso puede llevar imágenes que son gestionadas mediante la utilidad Raster Blaster que acompaña al programa. Se presenta como dos programas diferentes: las funciones de consulta y las de creación y administración de la base.

**Funciones**

- \* Diseño de la base de datos. Puede trabajar con todo tipo de documentos, con campos o sin ellos y éstos pueden estar o no indexados. El programa también ofrece muchas opciones para administrar la base, indexar o no los campos, excluirlos de las búsquedas e incluso crear una especie de tesoro. Además se pueden diseñar los formatos de consultas, tablas de contenido, editar el fichero de palabras vacías y proteger el acceso a documentos.

- \* Entrada de datos. El programa no proporciona ningún menú para la entrada de datos directa desde teclado, ya que presupone (como es tendencia en programas similares provenientes de EE.UU. o Gran Bretaña) que la información que gestiona proviene directamente de procesadores de textos, de otras bases de datos, o incluso por medios ópticos (OCR). Aún así el dis-



tribuidor del producto en España está elaborando un interfaz para poder introducir datos con teclado.

\* Búsqueda y recuperación de datos. Además del uso de los habituales operadores booleanos y truncamientos, WPL ofrece como novedad la recuperación de información calculando la probabilidad de pertinencia, es decir, que sólo da importancia en la búsqueda a los documentos que satisfacen más plenamente la demanda de información del usuario. Esta novedad hace que el tiempo de búsqueda se reduzca enormemente y que el éxito de recuperación sea más certero. Otra ventaja es la opción Expand, que sirve para calcular, a partir de los documentos existentes, qué términos están relacionados con el término o términos de búsqueda y luego los presenta en una ventana

como ideas para nuevos conceptos de búsqueda. La opción Concept realiza la misma función y además ejecuta la búsqueda de todos los documentos así vinculados y los ordena según su pertinencia.

**Hardware.** IBM PC compatible PC/PS con 286 o superior y Windows 3.x o superior.

**Proveedor.** Lo fabrica Personal Library Software y lo distribuye en España en castellano (incluyendo manual y tutorial): Compact Software Internacional. Villanueva, 8-4ª izda. 28001 Madrid. ☎ (91) 578 29 85.

**Precio.** Versión monopuesto: 180.000 pts. Versión para red: 850.000 pts.

#### OTROS PROGRAMAS

Además de las citadas, otras Bases de Datos Documentales son:

Freebase Personal (Base Libre), AskSam (Gofer-Ibérica), Micro-Questel (Atel), Hypersearch (Gould), Mistral (Bull), Status (Cap Gemini), Unidas (Untsys), etcétera.

#### BIBLIOGRAFIA

Esta información ha sido elaborada a partir de la documentación enviada por las empresas proveedoras y el apartado relativo a "Bases de datos documentales" del informe *Information Pack: On applying new technologies in Libraries. Greece and Spain*.

Como bibliografía básica, el libro más interesante editado en España es:

PASTOR, Oscar y BAQUEDANO, Arrate. *Bases de datos documentales: Cómo estructurar la información no estructurada*.- Valencia: Servei d'Assistència i Recursos Culturals, Diputació de Valencia, 1993.

\* Miguel A. Corcobedo y Belén Gómez son Diplomados en Biblioteconomía y Documentación.

## OTROS PROGRAMAS DE INTERÉS

#### HOMERO: GESTIÓN DE LIBRERÍAS

Programa de gestión comercial para librerías. Algunas de sus funciones son: *Mantenimiento* (búsquedas y fichas completas de libros, editoriales, colecciones, distribuidores, autores, materias, clientes y catálogos). *Gestión de compras* (borradores de compra, envío de pedidos, recepción de mercancías, generación de etiquetas, recepción de novedades, control de caducidades, gestión automatizada de stocks). *Gestión del proceso de ventas* (uso de Puntos de Venta Informatizados con posibilidad de integrar lectores ópticos y salidas de tickets de caja, control de notas de contado, albaranes y facturación, cartera de cobros). *Generación de recibos mensuales*, generación e impresión de códigos de barras, depuración de ficheros, apuntes contables con IVA, etcétera.

*Distribuye:* Atica Software. Mieres, 1. Entlo.B. 33208 Gijón. ☎ (98) 516 17 18. Fax: (98) 538 93 40.

#### BIBLIOS-HEMEROTECAS

Programa escrito en FoxPro (compatible con dBase) para gestión de los fondos de una hemeroteca (artículos de revista, periódicos y otros documentos). Comprende los módulos *Aniloc* o menú general, *Revista* o programa principal, *Hemeris* o suscripciones y *Herosfin* o mantenimiento, además de otras funciones de ayuda y vocabulario controlado.

Su virtud o característica más sobresaliente reside en la resolución, mediante los comandos de programación del software base, de los principales problemas que plantea la gestión automatizada de los fondos de una hemeroteca mediante un Sistema de Gestión de Bases de Datos Relacional.

*Para más información:* Julián Colina. [ver artículo publicado en *AABADOM, Boletín de la Asociación Asturiana de Bibliotecarios*, nº 2 (1993), pp. 6-11].

#### DOCUNAUTA: GESTIÓN DE ARCHIVOS

Permite gestionar archivos en papel, en microfilm, fototecas y videotecas. Combina una base de datos relacional con una documental.

El sistema se divide en dos módulos: *Génesis*: crea las distintas estructuras de referenciar la documentación, la pantalla necesaria para la entrada y visualización de los datos, tipo y longitud de los campos, obligatoriedad y tipo de recuperación de cada campo, etcétera. *Nauta*: realiza la gestión del archivo, mantenimiento de ficheros maestros, entrada de datos, consulta de información (operadores relacionales, documentales, lógicos y por rangos), gestión de espacio, control de préstamos, traslados y expurgos de documentación, control de usuarios, generador de informes, estadísticas.

*Distribuye:* Documenta Organización, S.A. General Martínez Campos, 15. 5º. 28010 Madrid. ☎ (91) 593 97 27. Fax: (91) 448 36 68.