

BIBLIOTECA
DE LA
UNIVERSIDAD DEL VALLE DE GUATEMALA

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ciencias y Humanidades

**MANEJO COMPUTARIZADO
DE
INFORMACION BIBLIOGRAFICA**


MANUEL EDUARDO CHINCHILLA SOLORZANO

**Modelo de trabajo profesional
presentado para optar al grado académico de
Licenciado en Ingeniería en Ciencias de la Computación**

**Guatemala
1992**

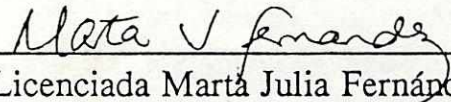
Vo. Bo.

(f)

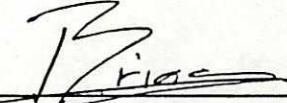

Licenciado David Alvarez
Asesor

Comité del Trabajo Profesional

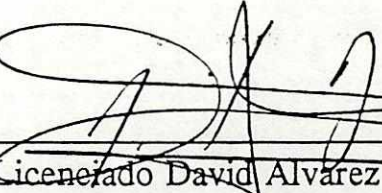
(f)


Licenciada Marta Julia Fernández

(f)


Licenciado Rodrigo Arias

(f)


Licenciado David Alvarez

Fecha de aprobación:

17 / JUNIO / 1972

CONTENIDO

	Páginas
PREFACIO	viii
RESUMEN	xi
I. INTRODUCCION	1
II. SISTEMAS DE INFORMACION BIBLIOGRAFICA	5
A. Estructura de un sistema de recuperación de documentos	8
B. Sistemas de recuperación de documentos en el contexto de un manejador de bases de datos	12
1. Modelo relacional	13
2. Modelo de red	15
3. Modelo jerárquico	15
C. Comparación entre sistemas de recuperación de documentos y manejadores de bases de datos	17
III. DESCRIPCION DEL SISTEMA dBIB	20
A. Visión global	20
1. Archivos y módulos del sistema	20
2. Interfase con el usuario	22
3. Manejo de errores	22
B. Características principales	26
1. Formato NLM de las referencias	26
2. Formato dBIB de las referencias	26
3. Palabras clave	29
4. Palabras del texto	30

5.	Diccionarios	30
6.	Selección y búsqueda de referencias	31
7.	Agrupación de referencias	32
8.	Mantenimiento de archivos	33
9.	Módulos y opciones	33
IV.	IMPLEMENTACION DEL SISTEMA	34
A.	Descripción general	34
B.	Archivos de dBIB	35
1.	MAIN: Datos generales de las referencias	35
2.	Archivo de textos	37
3.	Archivo de autores	37
4.	Archivo de revistas	38
5.	Archivos para las palabras clave y palabras del texto	38
6.	Otros archivos del sistema	39
C.	Descripción de las referencias	40
1.	Palabras clave	41
2.	Palabras del texto	44
D.	Asignación de códigos de identificación	46
E.	Módulos principales de dBIB	49
1.	Módulo "ADD"	49
2.	Módulo "FILE"	51
3.	Módulo "VIEW"	52
4.	Módulo "CRITERIA"	54
5.	Módulo "EDIT"	56
6.	Módulo "SORT"	56
7.	Módulo "PRINT"	56
8.	Módulo "OTHER"	56
V.	USO DEL SISTEMA BIBLIOGRAFICO dBIB	58

VI.	CONCLUSIONES Y RECOMENDACIONES	66
VII.	BIBLIOGRAFIA	68
	APENDICES	
A.	Estructuras de archivos de dBIB	70
B.	Manejo de errores inesperados en dBIB	76

LISTA DE FIGURAS

	Páginas
1. Estructura de archivos de STAIRS.	10
2. Relaciones en tercera forma normal.	14
3. Modelo de red.	16
4. Modelo jerárquico.	16
5. Elementos principales en un DBMS y un DRS.	18
6. Visión global del sistema dBIB.	21
7. Menú principal.	23
8. Selección de opciones en menús.	24
9. Mensajes en los menús.	25
10. Formato original de referencias enviadas por el NLM.	27
11. Listado de referencias producido por dBIB.	28
12. Módulos y submódulos de dBIB.	34
13. Archivos principales de dBIB.	36
14. Archivos para palabras clave en dBIB.	42
15. Palabras clave en dBIB.	44
16. Archivos para palabras del texto en dBIB.	44
17. Determinando el patrón para el campo IDO.	49
18. Matriz de referencias en la memoria.	53
19. Menú principal.	58
20. Referencias en "VIEW".	59
21. Formatos de referencias.	59
22. Referencias en formatos de una línea.	60
23. Menú de "CRITERIA".	60
24. Palabras clave.	61
25. Referencias con la palabra clave "TREATMENT".	61
26. Criterio de selección activado.	62
27. Menú de conectores lógicos.	62
28. Pantalla de autores en "CRITERIA"	63
29. Pantalla de "VIEW" en formato de una línea.	63

		Páginas
30.	Abstracto de una referencia.	64
31.	Final de sesión en dBIB.	64

PREFACIO

Científicos e investigadores en el mundo entero necesitan mantenerse actualizados en sus áreas de investigación. Para ello efectúan, principalmente, constantes lecturas de revistas y libros especializados. Los artículos en las revistas son la fuente de información más reciente sobre descubrimientos e investigaciones que se realizan en el mundo. Estas publicaciones periódicas también proporcionan al científico una cronología de los hallazgos en otras latitudes.

Las revistas científicas sobre salud se clasifican por temas, según la orientación de la publicación. Existen revistas como Parasitology, sobre parasitología; Journal of Tropical Medicine and Hygiene, que abarca la medicina tropical en general; Immunology Today, sobre inmunología; Journal of Infectious Diseases, referente a enfermedades infecciosas, etc. Para que un científico interesado en algún tema en especial se encuentre lo mejor informado posible, debe adquirir una gran cantidad de revistas, revisarlas, sacar fichas bibliográficas sobre los temas que le interesan y clasificarlas para usarlas en futuras consultas. Sólo en el área de salud existe un incontable número de revistas. El esfuerzo para mantenerse actualizado en estos tiempos resulta una tarea formidable para cualquier individuo.

La comunidad científica en los Estados Unidos está conciente de la importancia de mantener actualizados a sus investigadores. Los científicos en

el área de salud cuentan con la Biblioteca Nacional de Medicina en Bethesda, Maryland (National Library of Medicine o NLM). Gracias a convenios internacionales y debido a su importancia a nivel mundial, todos los artículos sobre medicina que son publicados en revistas, se remiten a esta biblioteca. De este modo se cuenta con toda la información en un solo lugar. En la biblioteca se clasifican e ingresan los artículos a un computador de acuerdo a un formato internacional para la publicación de artículos. Dicho formato se llama NLM por las siglas en inglés de la biblioteca.

La biblioteca cuenta con un departamento de consultas disponible al científico perteneciente a alguna de las organizaciones de salud acreditadas en la biblioteca. A través de este departamento, el científico puede obtener información sobre referencias bibliográficas, algún autor o tema en particular, etc. Las opciones para efectuar consultas son amplias y variadas. Los científicos de los Estados Unidos con autorización para el uso de este departamento ven resuelto el problema de mantenerse actualizado.

Dentro de la comunidad de científicos en salud de los Estados Unidos, existe un alto porcentaje que realiza su trabajo fuera de las fronteras de su país. La biblioteca NLM presta un servicio de correo aéreo a estos científicos. Envía mensualmente, a través de las embajadas de los Estados Unidos en diferentes países, listados de los artículos sobre temas que el científico le solicite. Optativamente, el científico puede obtener estos listados en diskette para usarlo en computadores personales compatible con IBM. Sin embargo, el científico no puede utilizar fuera de su país los servicios y facilidades del

departamento de consultas de la biblioteca en Bethesda. Lo único con que cuenta es con el listado de artículos.

Este problema ha llamado la atención del personal en las estaciones de campo fuera de los Estados Unidos. Medical Entomology Research and Training Unit/Guatemala (MERTU/G) es una estación de campo de la Oficina de Salud Pública de los Estados Unidos localizada en el campus de la Universidad del Valle de Guatemala. MERTU/G decidió apoyar el desarrollo de un sistema computarizado para el manejo de información bibliográfica. El objetivo del sistema es proporcionar a los científicos de esta estación, y a otros en situaciones similares, los recursos bibliográficos y facilidades para su manejo y así ayudarlos a mantenerse adecuadamente informados.

RESUMEN

En este trabajo se presenta un sistema computarizado para el manejo de información bibliográfica. El sistema se implementó en una computadora personal compatible con IBM, en un ambiente operativo DOS versión 5.

El proceso se divide en dos etapas:

Ingreso de los datos al sistema.

Las referencias bibliográficas deben estar almacenadas en código ASCII. El sistema organiza las referencias bibliográficas donde se distribuyen sus secciones, como: autores, abstracto, título, etc., en un sistema de archivos de bases de datos.

Manejo de la información.

Es la parte más poderosa del sistema, se encarga de realizar tareas, como: búsqueda, revisión, asociación y transporte de las referencias bibliográficas.

El sistema provee al usuario un conjunto de menús en pantallas que lo orientan e instruyen dentro del vasto número de opciones en cada menú.

El sistema se desarrolló siguiendo el concepto de bases de datos relacionales. Se normalizaron los archivos hasta llegar a su tercera forma normal. Cuando el sistema ingresa la información toma palabra por palabra, para luego asignarla a su archivo correspondiente y asociarla con el resto del sistema. Si la palabra no puede ser clasificada se almacena en un archivo especial que mantiene su asociación con todo el sistema. Se descartan las

palabras que no tienen significado, como: pronombres, artículos, adverbios, etc.

El usuario dispone de un grupo de opciones que le permiten realizar copias de las referencias bibliográficas, para ser intercambiadas con otros usuarios del sistema.

Para que el lector tenga un conocimiento básico sobre el sistema, también se presenta un historial de los estudios realizados y las soluciones planteadas.

I. INTRODUCCION

El problema de mantenerse adecuadamente informado de las últimas investigaciones y descubrimientos es especialmente crítico para la comunidad de científicos en misiones fuera de los Estados Unidos. El aprovechamiento de recursos ya existentes, como computadores personales en estas misiones, constituye una opción para resolver el problema.

El presente trabajo es una respuesta a las necesidades de información del grupo de científicos del Medical Entomology Research and Training Unit/Guatemala (MERTU/G). El propósito es proporcionarles un sistema computarizado para el manejo de información bibliográfica, aplicando conocimientos sobre bases de datos, sistemas de información y computadores personales.

En MERTU/G existen tres científicos investigadores extranjeros, así como varios científicos locales asociados que trabajan en forma conjunta. En esta unidad todos los científicos, extranjeros y locales, tienen una computadora a su disposición. Cada científico extranjero tiene el derecho de pedir a la biblioteca National Library of Medicine (NLM) en Bethesda, Maryland, copias de las referencias que él desee revisar. En el pedido se especifica el tema, por ejemplo: leishmaniasis, malaria y oncocercosis. Estas tres enfermedades tropicales son las que se estudian en esta unidad.

Anteriormente, el proceso para mantenerse bibliográficamente informado al día en MERTU/G no era uniforme. Cada científico decidía cómo manejar

los artículos o referencias bibliográficas. No existía un método único. En general el proceso era el siguiente: se revisa cada referencia y se separan las que no se consideran importantes. Aquéllas que llaman más la atención del científico se anotan manualmente en fichas o en un archivo en la computadora.

El tiempo que se requiere para revisar las referencias cada mes es considerable. El problema es mayor cuando el científico debe escribir reportes o consultar información. Para ello, el científico busca primero en sus fichas, luego en el primer grupo de referencias que seleccionó y por último en las que separó. Para cada consulta debe buscar, revisar, escoger y anotar la información que requiere, así una y otra vez hasta estar satisfecho. El número de referencias mensuales en un sólo tema (por ejemplo, leishmaniasis) es aproximadamente entre 50 y 75. El manejo manual de la información, así como el tiempo invertido en revisar y agrupar, son un gran inconveniente para este grupo de usuarios.

Con un sistema computarizado se persigue aprovechar las referencias bibliográficas enviadas en diskette y usarlas como entrada a dicho sistema. Una vez ingresadas, el usuario debe contar con opciones que le permitan agrupar, examinar, imprimir y cambiar (o modificar) el contenido de las referencias. Actualmente el sistema computarizado, llamado dBIB, se está usando en MERTU/G y también lo usan científicos en otras unidades.

dBIB reconoce el formato establecido por la biblioteca NLM, transfiere la información que ellos envían a un grupo de archivos que guardan las diferentes partes de las referencias y las relaciona a través de un código de

identificación que las distingue en forma única. La tarea del científico entonces se reduce a interactuar con un sistema computarizado sin tener que buscar física y manualmente entre los listados de referencias y/o fichas.

Además, dBIB ofrece opciones de selección de referencias fáciles de usar y, sobre todo, amigables al usuario. Permite seleccionar las referencias de acuerdo a sus atributos, tales como autor, palabras clave, palabras del texto, revista donde se publicó, etc. La selección puede incluir varios atributos, unidos entre sí a través de conectores lógicos, formando un criterio de selección. Al grupo de referencias que cumplan con dicho criterio se le puede asignar una palabra clave que identifique estas referencias permanentemente. Así en búsquedas subsecuentes, el usuario puede utilizar esta palabra clave en vez de volver a construir el criterio de selección.

dBIB ofrece opciones que permiten modificar el último criterio de selección, así como usar el último grupo de referencias encontrado como el universo para una nueva búsqueda. Las posibilidades de selección y agrupamiento de referencias son muy atractivas para el científico. Su información se vuelve dinámica, pudiéndose crear conjuntos de referencias bibliográficas a criterio del usuario, ya sea de manera permanente o temporal.

Las ventajas de este sistema sobre un proceso manual para manejar la información dan importancia y brindan interés al presente trabajo para la comunidad de científicos en MERTU/G. Este sistema fue desarrollado para funcionar en computadores personales compatibles con IBM con disco duro en un ambiente operativo DOS versión 5. Está desarrollado en el manejador de

bases de datos y lenguaje de programación CLIPPER¹ versión 5.01. Los archivos de datos son compatibles con el manejador de bases de datos DBASE III. Estas especificaciones aprovechan al máximo las condiciones existentes en MERTU/G. En las oficinas de salud de los Estados Unidos el manejador de bases de datos DBASE III es un standard. Los mensajes del sistema fueron escritos en inglés a petición de los científicos que apoyaron este esfuerzo.

¹ de Nantucket.

II. SISTEMAS DE INFORMACION BIBLIOGRAFICA

En el contexto de sistemas bibliográficos se identifican tres clases de sistemas de información [9]. Con el término *sistemas de información* se hace referencia a aquellos sistemas que facilitan el manejo de documentos, su almacenamiento y recuperación¹. Las tres clases de sistemas de información son: sistemas de recuperación de documentos o "Document Retrieval Systems" (DRS), manejadores de bases de datos o "Data Base Management Systems" (DBMS), y sistemas electrónicos de archivos o "Electronic Filing Systems" (EFS).

Como su nombre lo indica, los DRS se aplican para el manejo de documentos. Por su naturaleza, los documentos tienen datos poco estructurados. Es decir, son datos variables en formato libre. El formato libre, a diferencia del fijo, se refiere a textos escritos sin ninguna restricción de tamaño o gramática como mensajes, artículos, novelas, etc. Un caso específico de DRS es el manejo de información bibliográfica.

Los DRS generalmente manejan sólo un tipo de registros. Los registros suelen tener un número pequeño de campos, en su mayoría poco estructurados. A veces tienen un sólo campo, que contiene texto en formato libre.

En un DRS la recuperación de información se refiere a la selección de

¹ Generalmente el término *recuperación* equivale a la palabra "recovery". En este sentido se refiere a la recuperación de información dañada. Sin embargo en este trabajo *recuperación* se utiliza como traducción de la palabra "retrieval", refiriéndose al proceso de obtener información almacenada en el computador.

registros individuales cuyo contenido textual es sobre un t3pico en el cual est3 interesado el usuario. Lo cr3tico de este asunto es la manera efectiva en que se describe el t3pico a buscar y el contenido de los documentos. Los DRS proveen herramientas y facilidades para efectuar este proceso.

Una parte muy importante de los DRS es el 3ndice. El 3ndice es el conjunto de palabras clave¹ que identifican a los documentos del sistema. Estas palabras deben representar de la mejor manera posible el contenido de los documentos. El 3ndice de un DRS puede crearse manual o autom3ticamente.

Si dicho proceso es manual, el usuario especifica para cada documento sus respectivas palabras clave. Lo hace con base en su conocimiento y familiarizaci3n con el contenido del documento. Es de esperar que estas palabras correspondan, efectivamente, al contenido del documento. De ah3 que la recuperaci3n de documentos tienda a ser determin3stica. Sin embargo, no es totalmente determin3stica pues es pr3cticamente imposible especificar con exactitud el contenido del documento solamente a trav3s de las palabras clave que el usuario escoge [9].

En procesos autom3ticos, un algoritmo decide cu3les son las palabras clave del documento. Las palabras clave se escogen del texto del documento. Generalmente la selecci3n de palabras clave se basa en la frecuencia de aparici3n de las palabras en el documento. Un n3mero alto de apariciones suele indicar una mayor correspondencia entre esa palabra y el contenido del documento. En este caso, la recuperaci3n de informaci3n tiende a ser

¹ *Palabras clave* se usa como una traducci3n del ingl3s "keywords".

probabilística porque la palabra clave escogida por el algoritmo puede o no corresponder al contenido del documento.

Es conveniente eliminar del índice las variantes gramaticales para así reducir la redundancia. Algunos DRS están provistos de diferentes algoritmos que, en general, al encontrar una palabra clave antes de incluirla en el índice consultan con un diccionario de sinónimos. Si algún sinónimo ya fue asignado al documento entonces la palabra clave procesada no se incluye en el índice.

Los DBMS se aplican usualmente para manejar varios tipos de registros muy estructurados, o sea con formato fijo. Campos de textos con formato fijo usualmente son atributos, tales como nombre, dirección, ciudad, etc. La recuperación de información de los DBMS tiende a ser determinística, pues un registro es seleccionado sólo si un determinado campo contiene un valor específico [5].

Un manejador de bases de datos hace básicamente lo que su nombre implica: maneja datos. Más que sólo organizar registros, mantiene la información estructural que relaciona diferentes datos de información guardados en distintos archivos. El proceso de recuperación de la información depende del uso de relaciones establecidas entre los archivos de la base de datos.

Como ejemplo, una base de datos consta de los siguientes tres tipos de registros:

1. AUTOR, DIRECCION;
2. AUTOR, TITULO;
3. PALABRA CLAVE, TITULO, DOCUMENTO ó TEXTO;

Si se desea obtener las direcciones de los autores que escribieron los

documentos identificados por cierta palabra clave, se necesita de algún modo relacionar los tres tipos de registros. Debe indicarse de alguna forma que el campo TITULO del segundo registro está relacionado con el campo del mismo nombre del tercer registro y que el campo AUTOR de los dos primeros registros están relacionados entre sí.

Finalmente, el propósito de los EFS es guardar y recuperar información en archivos donde cada archivo es una colección de documentos relacionados [5]. Los EFS son muy adecuados para el manejo de información en oficinas, donde los registros son altamente estructurados y la información está organizada jerárquicamente en archivos, gavetas, cartapacios, grupos de hojas engrapadas, etc.

A. Estructura de un sistema de recuperación de documentos

Aunque los DRS principalmente se ocupan de recuperar información de registros de un sólo tipo (documentos), otros tipos de registros existen implícitamente en un DRS. Además del archivo de documentos existen archivos auxiliares que facilitan el manejo interno de la información, como por ejemplo el índice de las palabras clave y una lista de palabras sin significado como conectores, preposiciones, etc. Esta lista sirve cuando los DRS asignan automáticamente las palabras clave. Antes de asignar la palabra clave, los DRS buscan en la lista de palabras sin significado. Si la palabra aparece en esta lista, no se incluye en el índice. Es interesante notar que los DRS generalmente dejan revisar el índice de palabras clave y casi nunca permiten

revisar la lista de palabras sin significado.

El sistema más popular de DRS es el BRS. Este sistema se basa en STAIRS, un DRS desarrollado por IBM. Otros sistemas como ORBIT de System Development Corporation y MEDLINE del NLM presentan básicamente las mismas características que STAIRS [10]. A continuación se describe el sistema STAIRS.

Un documento en STAIRS contiene dos tipos de información:

1. el texto principal del documento, a partir del cual se construye el índice,
2. datos auxiliares que corresponden a los atributos del documento, tales como fecha de publicación, número de páginas, etc.

Los atributos del documento tienden a ser tipo numérico de formato fijo.

Puesto que rara vez un documento consiste de un sólo trozo continuo de texto, STAIRS tiene la habilidad de nombrar párrafos dentro del documento. El término *párrafo* en este contexto se refiere a las distintas secciones del documento, como el abstracto, los autores, el título, etc. STAIRS permite al usuario decidir en qué párrafos del documento debe realizarse la búsqueda.

La Figura 1 ilustra la estructura interna de STAIRS. La lista de palabras contiene todas las palabras en el texto, a excepción de las palabras sin significado. Para cada palabra en esta lista, el archivo invertido contiene el número de documentos que contienen esa palabra y el número total de veces que aparece la palabra en todos los documentos. Para cada vez que aparece la palabra, el archivo contiene la identificación del documento, el párrafo y la

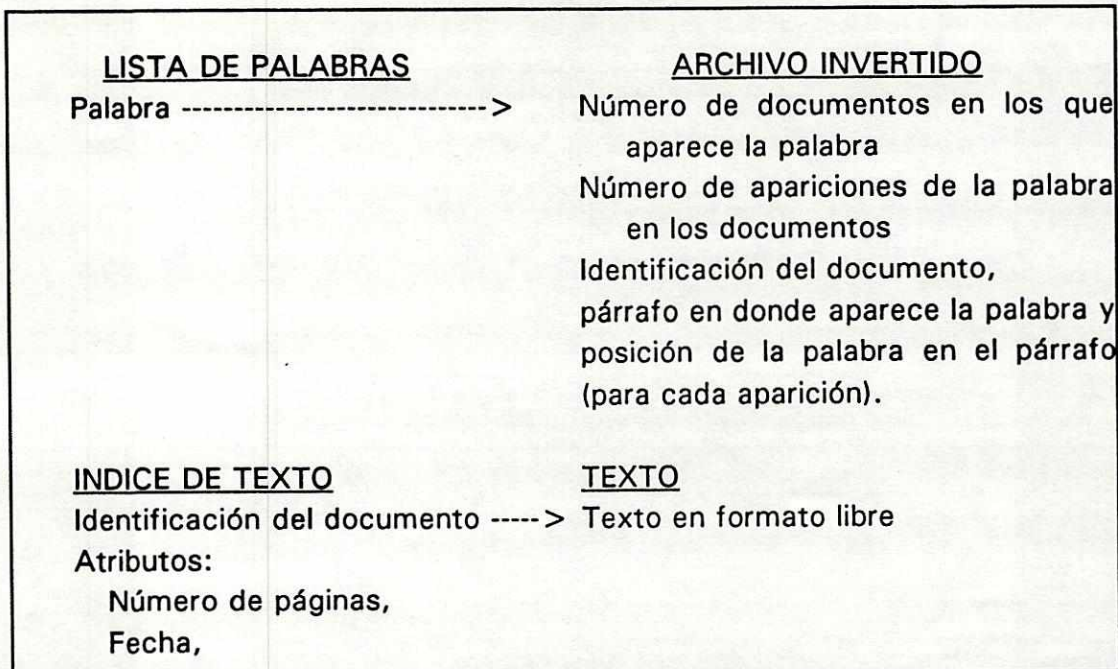


Figura 1. Estructuras de archivos de STAIRS.

posición dentro del párrafo donde aparece la palabra. El índice de texto contiene la identificación y atributos del documento.

Para las tres estructuras: la lista de palabras junto con el archivo invertido, el índice de texto y el texto, STAIRS ofrece tres comandos para la recuperación de documentos: SEARCH, que utiliza la lista de palabras; SELECT, que utiliza el índice de texto; y BROWSE, para ver el texto en su formato libre.

Cuando se usa SEARCH, se especifica una o varias palabras unidas por un conector lógico ("and", "or"), para formar expresiones de búsqueda. STAIRS asigna un número para cada expresión de búsqueda. Este número se puede usar en una expresión posterior para indicar a STAIRS que use el universo de documentos ya definido por dicha expresión. El símbolo "\$", al

final de una palabra, indica a STAIRS que la palabra está incompleta. El sistema buscará todas las palabras que empiecen con la serie de letras anteriores al símbolo "\$".

SEARCH	(entra al modo SEARCH)
1: INFORMATION AND RETRIEVAL	(busca documentos que contengan "information" y "retrieval")
2: 1 AND INTERACT\$	(en los documentos encontrados en el paso anterior, busca los que contengan palabras que comiencen con "interact")

En STAIRS se puede aprovechar la información sobre la posición de las palabras dentro del texto. Se pueden pedir los documentos que contengan dos palabras consecutivas, digamos "information retrieval", en el título.

3: INFORMATION.TITLE ADJ RETRIEVAL.TITLE

Si se deseara obtener los documentos con "information" o con "retrieval" entonces se escribe:

4: INFORMATION RETRIEVAL

Cuando no hay operador lógico entre las palabras, se considera un conector "or". Existe la opción RANK, la cual ordena los documentos encontrados de

acuerdo a la frecuencia de aparición de las palabras buscadas. Para buscar documentos con base en sus atributos (en el índice de texto), es necesario invocar el modo SELECT. Por ejemplo:

SELECT

5: 1 DATE > 1980

(inclirá los documentos encontrados en la primera expresión, más los documentos posteriores a 1980)

Después de haber seleccionado los documentos de interés con los comandos SEARCH y SELECT, el usuario puede verlos por medio del comando BROWSE.

MEDLINE contiene algunas características interesantes que no se encuentran en STAIRS. En particular utiliza un diccionario jerárquico que organiza las palabras clave (denominándolas MESH HEADINGS) de acuerdo a su amplitud. Para ello se designan ciertas palabras clave como un subconjunto de otras palabras clave. MESH HEADINGS se discutirán más a fondo en los capítulos siguientes.

B. Sistemas de recuperación de documentos en el contexto de manejadores de bases de datos

Existen tres modelos predominantes de bases de datos: el jerárquico, de red y el relacional. Respectivamente, éstos modelan el mundo como una

estructura de árbol, grafo y tabla. Al diseñar una base de datos, es necesario describir las relaciones entre los varios tipos de registros a manera que éstas se adecuen al modelo de bases de datos seleccionado.

A continuación se examinará cada uno de los tres modelos y se evaluará su funcionamiento y capacidad para implementar un DRS¹. Para ello se utilizará la misma base de datos bibliográfica. Los atributos de esta base de datos son: número o identificación del documento, título, autor, palabra clave, fecha y nombre de la revista. Conceptualmente se puede considerar que la base consiste de un solo registro, que a su vez consta de los seis campos anteriores. En la práctica, sin embargo, conviene separar los campos en diferentes tipos de registros de acuerdo al modelo en que se esté trabajando.

1. Modelo relacional. Intuitivamente una relación es una tabla sin orden específico de dos dimensiones, donde no hay dos filas iguales [4]. A la par del modelo relacional se desarrolló la teoría de normalización. Esta teoría es una disciplina para el diseño de relaciones con propiedades favorables para el proceso de actualización e integridad de datos. Para normalizar cualquier base de datos es fundamental el concepto de dependencia funcional. Para una relación R, el atributo B es funcionalmente dependiente del atributo A si, en todo momento, para cada valor de A en R no existe más de un valor B asociado en R. La teoría de normalización recomienda que una relación esté en su tercera forma normal. Una relación R

¹ Cabe mencionar que el modelo relacional y el de red ya han sido usados como base para un DRS [5,6,9].

está en tercera forma normal si es una tabla y para cada conjunto C de atributos de R. Existe algún atributo de R que no esté en C y que sea funcionalmente dependiente de C, entonces todos los atributos en R son funcionalmente dependientes de C. Intuitivamente se puede pensar que cada relación en tercera forma normal describe un único concepto o entidad. La Figura 2 muestra la base de datos del ejemplo como tres relaciones en su tercera forma normal.

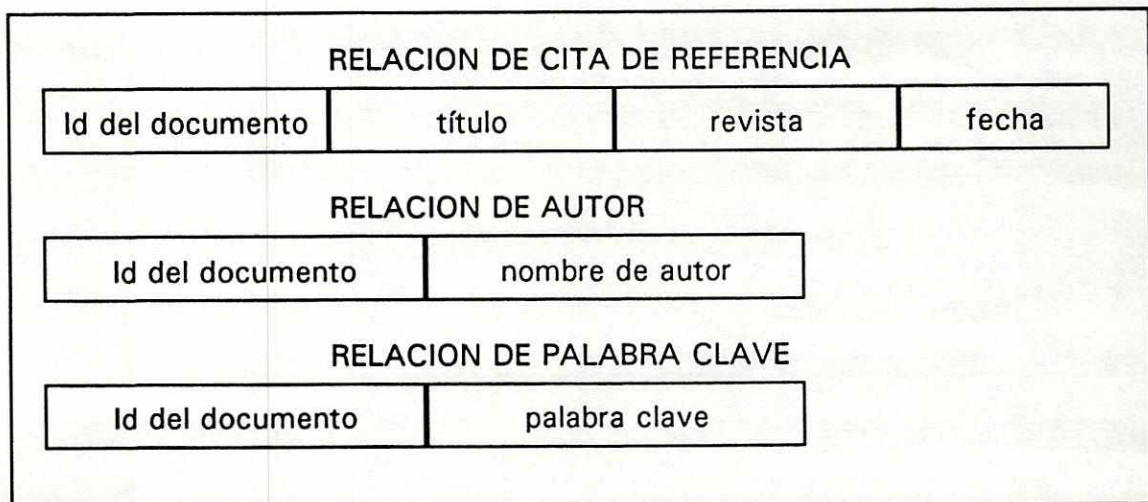


Figura 2. Relaciones en tercera forma normal.

Las ventajas que ofrece el modelo relacional para la recuperación de información son:

1. **Simplicidad del modelo:** El usuario puede ver la estructura y relaciones del sistema en una manera clara, simple y coherente.
2. **Base teórica:** Una teoría muy fuerte respalda este modelo, fomentando el desarrollo de muchos DBMS basados en este modelo.
3. **Independencia de datos:** El usuario no interviene en los detalles de

almacenamiento y recuperación de información.

4. Simetría en la recuperación de información: Por ejemplo, se usa el mismo algoritmo para encontrar todos los documentos de un autor y todos los autores de un documento.

2. **Modelo de red.** Las especificaciones para el modelo de red fueron publicadas en el Reporte del Grupo DATA BASE TASK (DBTG) de la Conferencia sobre Lenguajes y Sistemas de Datos (CODASYL) [4]. En el presente trabajo se describe el modelo de red de acuerdo a este reporte.

En el modelo DBTG la información está representada por registros y cadenas. Los registros representan entidades y definen las estructuras de las tablas. Las cadenas son usadas para especificar las relaciones o conexiones entre los distintos registros o entidades. La Figura 3 muestra un grafo que representa los registros y sus relaciones para el ejemplo bibliográfico.

En este ejemplo se manejan relaciones de varios a varios. En el caso de autores y documentos se introduce un registro intermedio llamado *documento/autor* que sirve para relacionar a los documentos con sus autores y viceversa. Lo mismo sucede en la relación entre palabras clave y documentos. El proceso de manejar el registro intermedio agrega cierto grado de complejidad al sistema.

3. **Modelo jerárquico.** El modelo jerárquico es un caso particular del modelo de red. Los arcos o flechas en el diagrama de estructura para este modelo deben formar un árbol ordenado. El modelo jerárquico para el ejemplo bibliográfico se muestra en la Figura 4.

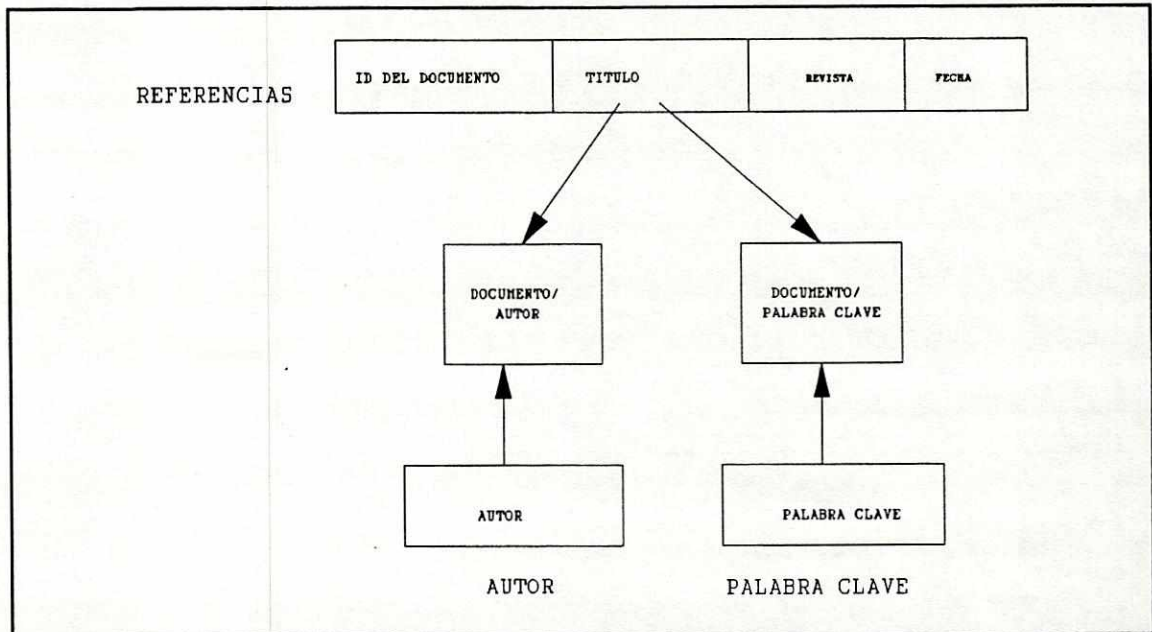


Figura 3. Modelo de red.

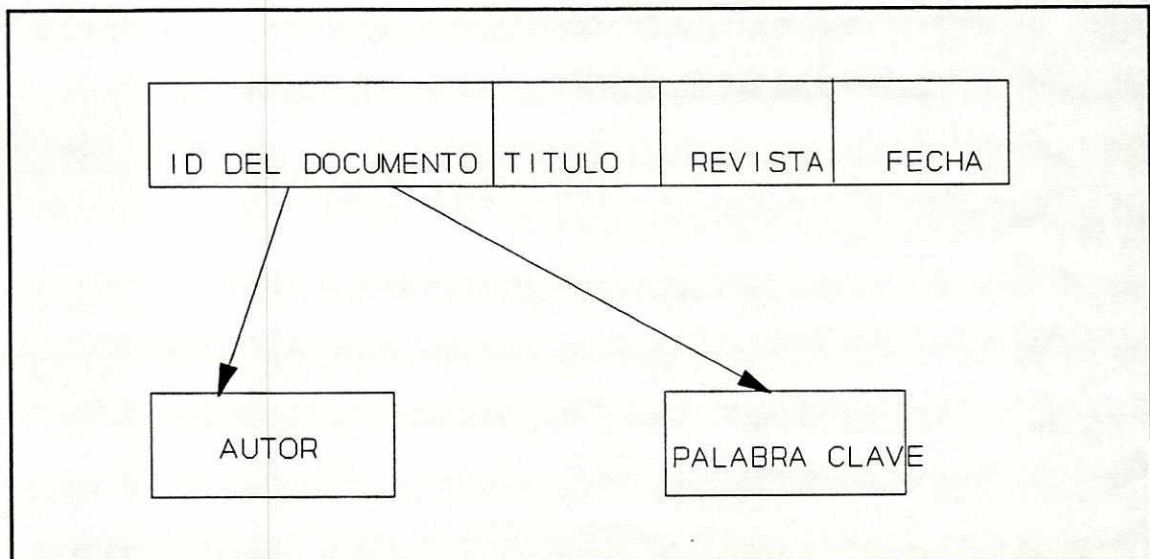


Figura 4. Modelo jerárquico.

En este modelo se presentan muchas dificultades. Por ejemplo, es fácil obtener los autores de algún documento. Sin embargo, obtener los documentos

de algún autor es muy difícil bajo este modelo. La jerarquía impuesta impide la simetría en la recuperación de información, pues las relaciones sólo van en un sentido, de ahí el término *jerárquico*.

C. Comparación entre manejadores de bases de datos y sistemas de recuperación de documentos

Los DBMS se basan en uno de los tres modelos que se mencionaron anteriormente. Los DRS se basan en modelos diseñados específicamente para la recuperación de documentos. Los DBMS son en general sistemas más complejos que los DRS. En los DBMS se pueden manejar diferentes tipos de registros mientras que en los DRS generalmente sólo existe una clase de registro.

Las referencias bibliográficas incluyen el texto de los abstractos así como el nombre del autor, volumen, número de páginas, fecha de publicación, etc. Aunque los DBMS manejan atributos con formato bien definido y los DRS procesan específicamente textos, la unión de ambos sistemas puede producir aplicaciones completas y eficientes para el manejo de información bibliográfica. El problema de integrarlos para manejar la información bibliográfica es básico para el desarrollo del presente trabajo.

Un punto esencial de notar en STAIRS, y en general en todos los DRS, es que diferencian marcadamente entre el texto y los atributos de los documentos. Cada uno se almacena y maneja de manera diferente. Esta falta de integración de la información ha incidido en que los DBMS y DRS se hayan desarrollado separadamente a través de la historia. En la Figura 5 se enumeran

los componentes principales de un DBMS y los de un DRS.

DBMS	DRS
Grupos de estructura de datos (varios archivos con diferentes estructuras).	Una sola estructura de datos (textos).
Un esquema que describe la organización de la base de datos.	No hay esquema.
Rutinas disponibles en lenguajes de alto nivel para el manejo de la base de datos	No hay rutinas externas para el manejo del DRS. En algunos casos sólo existen rutinas para reconstruir o formar los índices usados en los DRS
Un lenguaje para interactuar con el usuario y manejar o transformar la información	Un lenguaje muy sencillo y limitado para interactuar con el usuario sin permitir el manejo y transformación de los documentos

Figura 5. Elementos principales en un DBMS y un DRS.

Para combinar la funcionalidad de estos dos sistemas para el desarrollo de un sistema de información bibliográfica se debe trabajar sobre una plataforma que ataque las debilidades y aproveche las ventajas de cada uno. El modelo relacional de los DBMS es más adecuado para el desarrollo de una aplicación DRS que los otros modelos, principalmente porque permite una mayor versatilidad en la recuperación de información. Además, en el mercado de computadores personales existen muchos paquetes manejadores de bases de datos que se fundamentan en este modelo. Estos paquetes ofrecen en su mayoría lenguajes de programación por medio de los cuales se pueden crear rutinas que ejecuten las funciones encontradas en los DRS y STAIRS. Se puede también crear interfases con el usuario más amigables que las tradicionalmente usadas en los DRS.

CLIPPER no es sólo un manejador de bases de datos relacional. Su lenguaje de programación permite crear rutinas que simulan las funciones de los DRS. Los archivos de índices, como los usados en STAIRS, se pueden crear y manejar fácilmente desde CLIPPER. Además esta herramienta posee rutinas predefinidas para la definición y uso de pantallas, lo que permite que la interacción entre el sistema y el usuario sea fácil y conocida.

III. DESCRIPCION DEL SISTEMA dBIB

dBIB es un sistema de información bibliográfica para computadores personales. Maneja referencias bibliográficas según el formato de la biblioteca NLM. Se basó en los ejemplos de DRS ya existentes como STAIRS y MEDLINE. Ofrece las principales características de dichos DRS con una interfase con el usuario amigable y poderosa. Además ofrece: una gran variedad de formatos para ver o imprimir las referencias, menús que guían al usuario en la formación de criterios de selección, revisión y creación de palabras clave, adición de palabras clave a una referencia o un grupo de referencias, adición de comentarios propios del usuario a las referencias, rutinas de mantenimiento y respaldo de los archivos (estas últimas permiten el intercambio de información entre usuarios de dBIB). Todo ello hace del sistema una herramienta útil y fácil de emplear en la tarea de organización y búsqueda de referencias bibliográficas.

El sistema dBIB consta de un módulo general y ocho submódulos principales que controlan todas las opciones que ofrece el sistema. Cada submódulo se subdivide en partes más pequeñas que controlan acciones más específicas.

A. Visión global

- 1. Archivos y módulos del sistema.** La Figura 6 muestra gráficamente el sistema en forma global. Se muestran los nombres de los archivos

y módulos más importantes que usa el sistema dBIB.

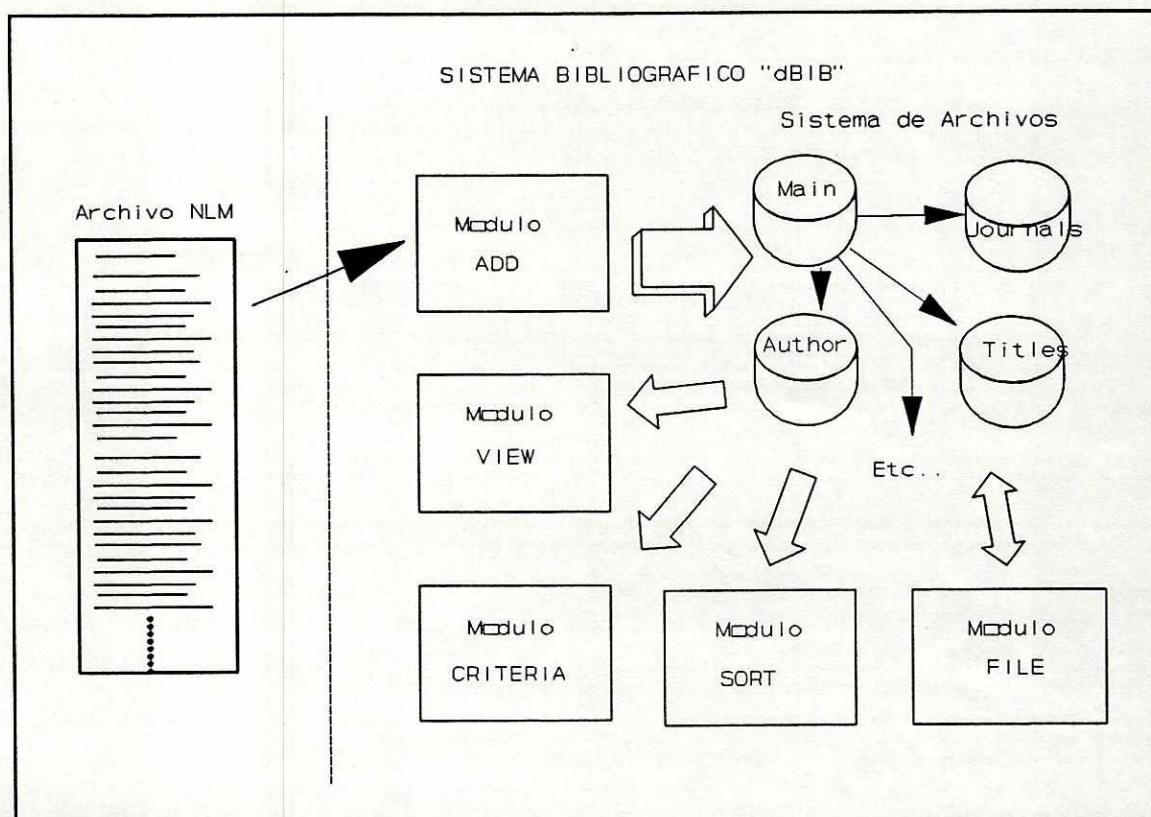


Figura 6. Visión global del sistema dBIB.

La flecha que va del archivo NLM al grupo de archivos del sistema representa el traslado de las referencias en su forma original al sistema dBIB. El módulo "ADD" realiza este traslado y distribuye los párrafos de las referencias entre el grupo de archivos MAIN, JOURNALS, TITLES, etc., de dBIB. El archivo MAIN contiene la identificación y datos generales de la referencia y se relaciona con los demás archivos donde se encuentra el resto de la información de las referencias. En TITLES se guardan los títulos y abstractos de las referencias, en JOURNALS los nombres de las revistas donde

aparecieron las referencias y en AUTHOR sus autores. El módulo "VIEW" se utiliza para presentar las referencias en variados formatos, "CRITERIA" ayuda al usuario en la formación de un criterio para seleccionar referencias, "SORT" ordena las referencias (por ejemplo, por fecha de publicación o por orden alfabético de nombre de autor) y "FILE" maneja las operaciones de archivos, como copia, eliminación y copias de respaldo de las bibliografías.

2. Interfase con el usuario. Al iniciarse el sistema, se presenta al usuario un menú principal (ver Figura 7). En él se puede elegir, con las teclas de flechas ($\uparrow\downarrow$), cualquiera de ocho posibles opciones correspondientes a cada uno de los módulos principales del sistema.

La primera línea en la Figura 7 es la línea de status. Esta línea indica cuántas referencias hay en la bibliografía¹ que se está operando, si está ordenado por algún campo o no y el criterio actual de selección de las referencias. El menú presenta los nombres de las opciones en mayúsculas y una breve explicación de su función a la derecha de la opción. En la parte inferior derecha del marco del menú se despliega el nombre de la bibliografía que se está revisando. La línea inferior de la pantalla indica la función de algunas teclas para este menú. Todos los otros menús del sistema se parecen al de la Figura 7, lo cual hace más consistente la interacción del usuario con las diferentes partes del sistema.

En todos los menús, las cuatro teclas de las flechas ($\uparrow\downarrow\rightarrow\leftarrow$) se usan

¹ Al conjunto de archivos del sistema que contienen la información de las referencias se le llama *bibliografía*.

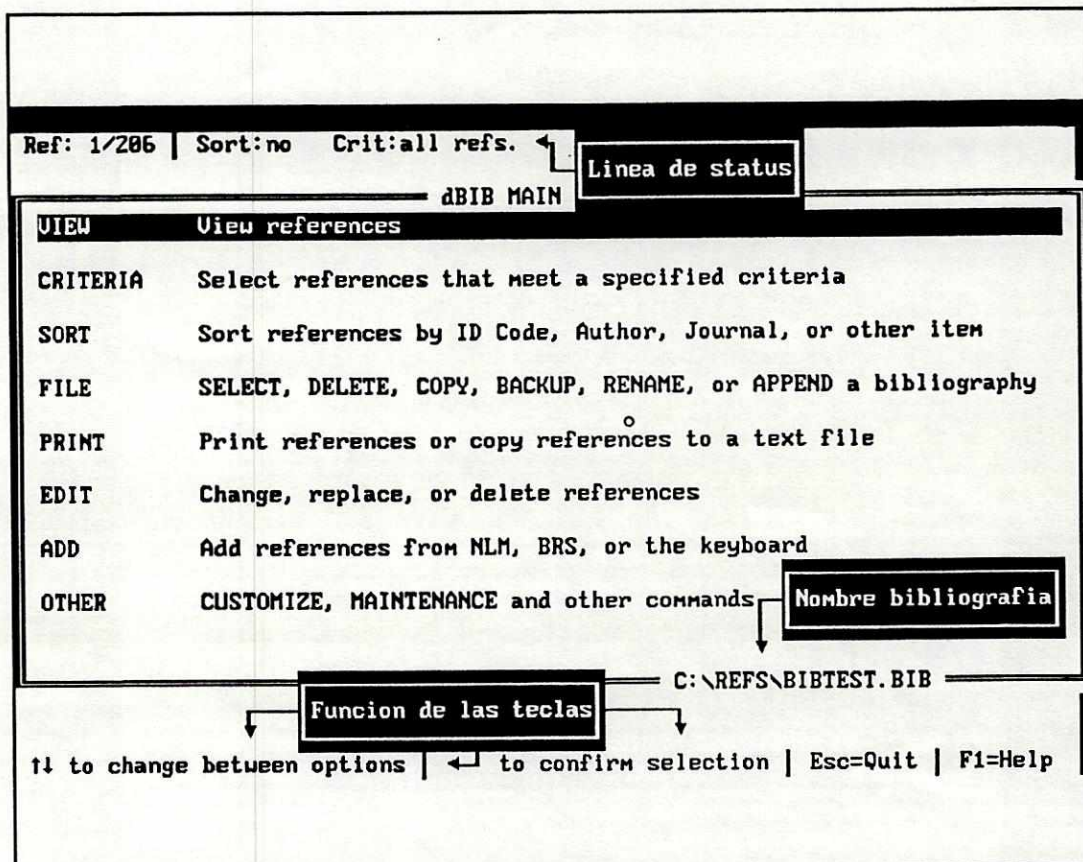


Figura 7. Menú principal.

para seleccionar las opciones. En la Figura 8 aparece la pantalla que se presenta al seleccionar "VIEW" en el menú principal. Las opciones "DISPLAY", "JUMP", "EDIT", etc., debajo de la línea de status, constituyen el menú de este módulo y pueden seleccionarse usando las teclas de las flechas izquierda y derecha (→ ←). Algunas de estas opciones contienen a su vez subopciones, como en el caso de "JUMP" (ver en la figura). Las subopciones se seleccionan a través de las teclas de flechas arriba y abajo (↓ ↑). Por el contrario la opción "DISPLAY" no contiene subopciones y las flechas (↓ ↑) se usan para ir de una o otra referencia.

```

Ref: 4/206 | Sort:no | Crit:all refs.
DISPLAY  JUMP  EDIT  PRINT  MARK  FORMAT  Q IT

```

Mesa-Ualle	GO TO	action of platinum (I) and platinum (IU) complexes on	01
	FIND	telf	
	TOP		
Albanese G	END	leishmaniasis.	traconazole.
	NEXT	atol	1989 125(11):1540-2 LE9001.002
	PREVIOUS		
Pratlong F		eous presence in dogs of 2 zymodemes of the Leishmania	
		Ann Parasitol Hum Comp	1989 64(4):31-4 LE9001.003
Dourado MI		[Epidemiology of tegumentary American leishm niasis and its rel	
		Rev Saude Publica	
Gomes A de		Ecologic aspects of American teg	
		Rev Inst Med Trop Sao Paulo	1989 31(1):324-1 LE9001.005
Cairns BR		Developmentally the genetics of leishmania encodes a putative	
		Proc Natl Acad Sci U S A	1989 86(70):7682-6 LE9001.006

C:\REFS\BIBTEST.BIB

```

+<->=change menu option. |↑↓=move between Move options|*=marked ref|ESC=Quit
Go to a desired number of reference

```

Figura 8. Selección de opciones en menús.

El sistema generalmente despliega mensajes adicionales que indican al usuario los pasos necesarios para realizar determinada acción. La Figura 9 muestra los mensajes que aparecen al seleccionar "AUTHOR" en el menú de "CRITERIA".

La opción seleccionada siempre se ilumina o realza. En monitores a color el usuario puede decidir los colores que indiquen el realce de la selección.

La tecla *escape* sirve para abandonar el menú actual donde se encuentra el usuario o para cancelar la última acción dentro de algún proceso.

Todos los procesos piden al usuario una confirmación final antes de ejecutar cualquier acción.

En algunas pantallas donde se pide al usuario elegir entre diferentes

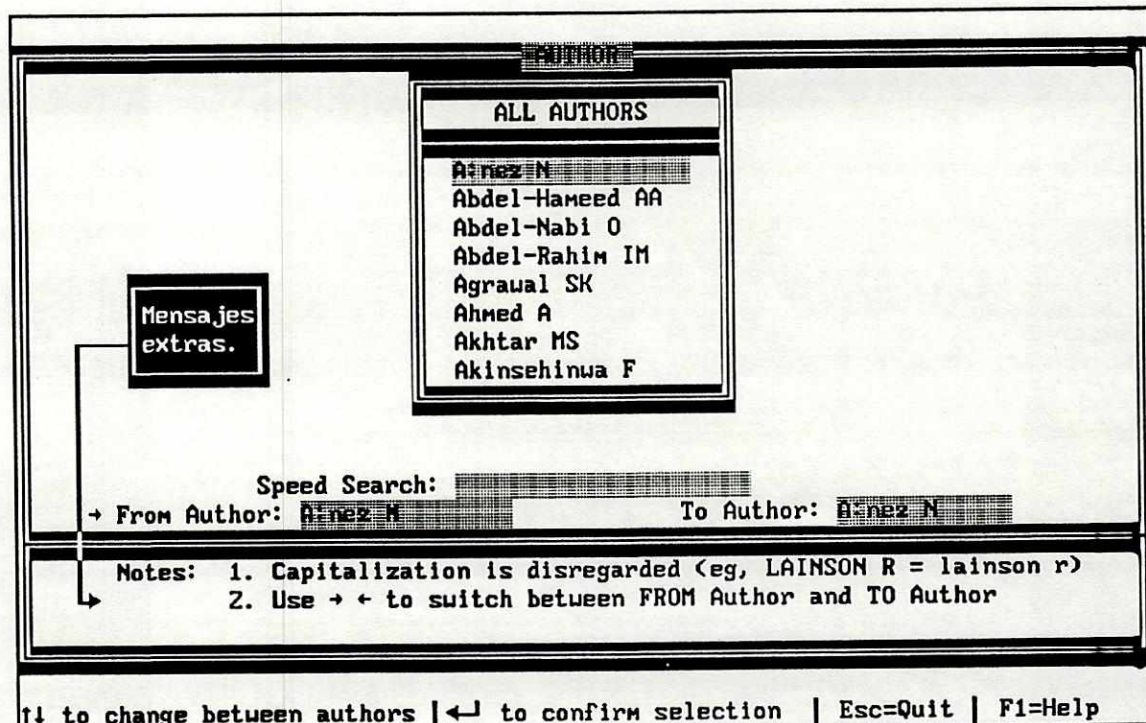


Figura 9. Mensajes en los menús.

opciones, existe una opción ya preseleccionada. El usuario puede cambiarla o aceptarla. El usuario puede cambiar estas opciones predefinidas en forma permanente de acuerdo a sus preferencias.

3. Manejo de errores. El sistema incluye un proceso para el manejo de errores imprevistos, el cual evita que el programa se detenga de manera anormal. Este proceso indica el programa y la línea donde ocurrió el error, que es una información importante para su corrección. Errores imprevistos se pueden dar en las siguientes situaciones: el impresor no está listo o no tiene hojas, falta espacio libre en el disco duro, hay conflictos en la memoria principal, etc. Una descripción detallada de este proceso se incluye en el Apéndice B.

B. Características principales

1. **Formato NLM de las referencias.** En la Figura 10 se muestra un ejemplo de las referencias en su formato original enviadas por el NLM. El usuario debe decidir qué referencias serán parte de qué bibliografía. El elemento individual dentro una bibliografía es la referencia. Cada referencia tiene un número correlativo que las diferencia dentro del archivo enviado por el NLM. La referencia consta de varios párrafos: título (TI), autores (AU), abstracto (AB), lenguaje (LA), "Mesh Headings"¹ (MH), identificación única (UI), dirección (AD), revista (SO) y fecha de publicación (DP). En ocasiones pueden existir otros párrafos que dBIB "no reconoce". En ese caso dBIB almacena todos los párrafos desconocidos en un archivo diferente conservando las siglas que los identifican.

Las referencias en su formato original se ingresan al sistema de archivos de dBIB a través del módulo "ADD". La información de cada referencia se guarda en archivos separados que se relacionan entre sí, según se discute en el siguiente capítulo.

2. **Formato dBIB de las referencias.** La Figura 11 muestra un listado generado por dBIB, en el cual la información de la referencia ya ha sido distribuida en el sistema de archivos de dBIB. La identificación de la primera referencia es LE9001.001, y LE9001.002 la de la segunda. Estos códigos de identificación los proporciona el usuario.

¹ "Mesh Headings" se deriva del término en inglés "Medical Subheadings". El NLM asigna palabras clave a las referencias y se incluyen en el párrafo "Mesh Headings".

prt au, ti, tt, ab, la, mh, ui, dp, ad, so compr

PROG:

1

AU - Mesa-Valle CM ; Craciunescu D ; Parrondo-Iglesias E ; Osuna A
TI - In vitro action of platinum (II) and platinum (IV) complexes on trypanosoma cruzi and Leishmania donovani.
AB - The in vitro activities of 22 neutral Pt(II) complexes, 4 Pt(IV) complex salts, and one Pt(II) complex salt on epimastigote forms of Trypanosoma cruzi and promastigote forms of Leishmania donovani were studied. Only 2 out of the 18 complexes with cis-Pt(DDH)_n structure completely inhibited the growth of the epimastigote forms of T. cruzi and the promastigote of L. donovani. Against T. cruzi the cis-Pt(L)_n(Cl)₂ complexes showed no activities, but the complex in which the (L) ligand was stilbamidine did show a limited activity against L. donovani. Of the Pt(IV) complex salts, [Pt X₆]H₂(L) were very active against the epimastigote forms of T. cruzi. However, the Pt(II) complex salt [cis-Pt Cl₄]H₂ (pentamidine) showed no activity. The actions of the Pt(IV) complex salts on the promastigote forms of L. donovani were different and only those salts in which the (L) ligand was pentamidine or stilbamidine induced parasite growth inhibition. The one Pt(II) salt tested showed no antiparasitic activity. At the same time, the known antiparasitic activities of the nifurtimox and pentamidine molecules were confirmed.
LA - Eng
MH - Animal ; Leishmania donovani/*DRUG EFFECTS/GROWTH & DEVELOPMENT ; Organoplatinum Compounds/*PHARMACOLOGY ; *Trypanocidal Agents ; Trypanosoma Cruzi/*DRUG EFFECTS/GROWTH & DEVELOPMENT
UI - 90056664
DP - 1989 Aug
AD - Department of Parasitology, Faculty of Science, University of Granada, Spain.
SO - Arzneimittelforschung 1989 Aug;39(8):838-42
CONTINUE PRINTING? (YES/NO)

USER:

Y

PROG:

2

AU - Albanese G ; Giorgetti P ; Santagostino L ; Crippa D ; Sala G
TI - Cutaneous leishmaniasis. Treatment with itraconazole.
AB - The chemotherapy of leishmaniasis is still far from satisfactory. Itraconazole, one of the most recent azole antifungal agents, appears to be well tolerated in man. Two patients with cutaneous leishmaniasis (Leishmania tropica), contracted in Saudi Arabia, were treated orally for 2 months with itraconazole (100 mg/d). Both recovered without side effects or abnormalities in their main biologic parameters.
LA - Eng
MH - Adult ; Antifungal Agents/*THERAPEUTIC USE ; Case Report ; Follow-Up Studies ; Human ; Ketoconazole ; Leishmaniasis/*DRUG THERAPY ; Male
UI - 90055236
DP - 1989 Nov
AD - Department of Dermatology, S. Gerardo Hospital, Monza, Italy.
SO - Arch Dermatol 1989 Nov;125(11):1540-2

Figura 10. Formato original de referencias enviadas por el NLM.

LE9001.001

Mesa-Valle CM, Craciunescu D Parrondo-Iglesias E, Osuna A. In vitro action of platinum (II) and platinum (IV) complexes on trypanosoma cruzi and Leishmania donovani. *Arzneimittelforschung* 1989;39(8):838-42.

The in vitro activities of 22 neutral Pt(II) complexes, 4 Pt(IV) complex salts, and one Pt(II) complex salt on epimastigote forms of Trypanosoma cruzi and promastigote forms of Leishmania donovani were studied. Only 2 out of the 18 complexes with cis-Pt(DDH)_n structure completely inhibited the growth of the epimastigote forms of T. cruzi and the promastigote of L. donovani. Against T. cruzi the cis-Pt(L)_n(Cl)₂ complexes showed no activities, but the complex in which the (L) ligand was stilbamidine did show a limited activity against L. donovani. Of the Pt(IV) complex salts, [Pt X₆]H₂(L) were very active against the epimastigote forms of T. cruzi. However, the Pt(II) complex salt [cis-Pt Cl₄]H₂ (pentamidine) showed no activity. The actions of the Pt(IV) complex salts on the promastigote forms of L. donovani were different and only those salts in which the (L) ligand was pentamidine or stilbamidine induced parasite growth inhibition.

Keywords

SUBHEADING; GROWTH & DEVELOPMENT
TAXONOMY & CLASSIFICATION; L. DONOVANI
TREATMENT; DRUG EFFECTS

Mesh Headings

Animal ; Leishmania donovani/*DRUG EFFECTS/GROWTH & DEVELOPMENT ; Organoplatinum Compounds/*PHARMACOLOGY ; *Trypanocidal Agents ; Trypanosoma Cruzi/*DRUG EFFECTS/GROWTH & DEVELOPMENT.

Data

Reprint : N
Journal/Book : Journal
Language : Eng
Marked : No
Location : GLUCANTIME
Unique ID : 90056664
DP : 1989 Aug
AD : Department of Parasitology, Faculty of Science, University of Granada, Spain.

LE9001.002

Albanese G, Giorgetti P, Santagostino L, Crippa D, Sala G. Cutaneous leishmaniasis. Treatment with itraconazole. *Arch Dermatol* 1989;125(11):1540-2. The chemotherapy of leishmaniasis is still far from satisfactory. Itraconazole, one of the most recent azole antifungal agents, appears to be well tolerated in man. Two patients with cutaneous leishmaniasis (*Leishmania tropica*), contracted in Saudi Arabia, were treated orally for 2 months with itraconazole (100 mg/d). Both recovered without side effects or abnormalities in their main biologic parameters.

Keywords

DISEASE SYNDROMES; CUTANEOUS LEISHMANIASIS
RESTRICTOR; ADULT
RESTRICTOR; HUMAN
RESTRICTOR; MALE
TAXONOMY & CLASSIFICATION; L. TROPICA
TREATMENT; ANTIFUNGAL AGENTS
TREATMENT; DRUG THERAPY
TREATMENT; TREATMENT

Mesh Headings

Adult ; Antifungal Agents/*THERAPEUTIC USE ; Case Report ; Follow-Up Studies ; Human ; Ketoconazole ; Leishmaniasis/*DRUG THERAPY ; Male.

Data

Reprint : N
Journal/Book : Journal
Language : Eng
Marked : No
Location : GLUCANTIME
Unique ID : 90055236
DP : 1989 Nov
AD : Department of Dermatology, S. Gerardo Hospital, Monza, Italy.

Figura 11. Listado de referencias producido por dBIB.

A continuación de la identificación de la referencia se encuentran los nombres de los autores, el título del artículo, el nombre de la revista, fecha, volumen, páginas y abstracto del artículo.

Las palabras clave que dBIB tiene asignadas a esa referencia se muestran bajo el título "Keywords" en la Figura 11. Bajo el nombre de "Mesh Headings" se encuentran las palabras clave que el NLM asignó a la referencia. En ocasiones, el grupo de palabras clave bajo "Keywords" difiere del grupo en "Mesh Headings". Esto sucede porque el usuario decidió descartar algunas palabras clave del NLM.

Por último, en la Figura 11 se reporta bajo el nombre "Data": si el usuario tiene en su poder reimpresiones¹ del artículo, si el artículo es parte de un libro o revista, el lenguaje original en que fue escrito, si la referencia está marcada o no. El usuario marca una referencia para distinguirla de otras referencias y luego efectuar alguna operación (eliminación, copia, adición de palabras clave, etc.) sólo con las marcadas. En "Data" también se lista la localización de la referencia (por ejemplo, en un folder con el título de Glucantime), el identificador único del NLM para esta referencia, fecha de publicación y la dirección del lugar donde se escribió la referencia.

3. Palabras clave. dBIB asocia palabras clave a cada referencia. Las palabras clave son palabras que dicen de qué se trata el contenido de la referencia. dBIB usa las palabras clave para el proceso de selección y búsqueda de referencias que se explica en una sección más adelante.

¹. La copia del artículo como fue publicado en la revista se llama *reimpresión*.

dBIB proporciona al usuario una serie de rutinas que le ayudan a crear un diccionario de palabras clave para la bibliografía. Esta lista la utiliza dBIB para asignar las palabras clave a cada referencia de la bibliografía. La manera en que se realiza esta asignación se describe en el capítulo de implementación.

Para formar el diccionario de palabras clave, dBIB toma las palabras que aparecen en los párrafos de "Mesh Headings" en el grupo de referencias con formato NLM. De esta lista preliminar, el usuario decide cuáles serán las palabras clave a utilizar en la bibliografía. Si el usuario desea agregar sus propias palabras clave al diccionario, lo puede hacer.

4. Palabras del texto. dBIB crea automáticamente un diccionario de palabras del texto para el texto en formato libre de una referencia. Las palabras las toma del texto y excluye las que no tienen significado (preposiciones, artículos, adverbios, etc.). El objetivo del diccionario de palabras del texto es ofrecer al usuario novato en la materia, que desconoce los términos usados en las palabras clave, una alternativa para especificar las referencias a ser buscadas.

5. Diccionarios. Los diccionarios de dBIB los utiliza el usuario para asistirlo en la formación del criterio de selección de referencias. Los cuatro diccionarios principales en dBIB son: el de autores, el de revistas, el de palabras clave y el de palabras de texto. El diccionario de autores, revistas y palabras del texto se forman automáticamente de las referencias del NLM mientras que el diccionario de palabras clave lo define manualmente el usuario. Estos pueden ser modificados por el usuario

6. Selección y búsqueda de referencias. dBIB ofrece dos formas principales para realizar búsquedas de referencias en las bibliografías. Una es determinística pues se buscan referencias con determinados valores en sus campos. Por ejemplo, se buscan las referencias pertenecientes al Doctor Byron Arana, es decir el campo *autor* que contenga dicho nombre. La otra búsqueda es probabilística donde el usuario, a través de las palabras clave y/o las palabras del texto, intenta recuperar las referencias que traten sobre determinado tópico.

La selección de referencias se realiza a través del módulo "CRITERIA", que asiste al usuario en la formación de su criterio de selección. Este criterio puede estar formado por campos determinísticos (autor, revista, año de publicación, etc.) y/o campos probabilísticos (palabras clave y palabras del texto), unidos por conectores lógicos y agrupados con paréntesis para cambiar la precedencia de evaluación en la expresión lógica.

Cuando en el criterio se usan campos que tienen diccionarios en dBIB, "CRITERIA" despliega el diccionario en la pantalla para que el usuario elija entre los valores válidos que existen en la bibliografía. Para que el usuario localice rápidamente la palabra que le interesa en el diccionario, dBIB lo muestra ordenado alfabéticamente. Además, dBIB va buscando la palabra letra por letra según la vaya escribiendo el usuario. Va señalando, dónde se encuentra dicha palabra parcial en el diccionario, así hasta encontrarla o descubrir que no existe en la bibliografía. Cuando el usuario escribe una letra de la palabra que no corresponde a ninguna palabra en el diccionario, dBIB se

lo advertirá con un mensaje de alerta. Entonces permitirá al usuario continuar con la última letra válida o cancelar la operación.

7. Agrupación de referencias. dBIB permite manejar las referencias en forma individual o en grupos. El resultado de una selección exitosa es siempre un grupo de referencias. A través de "VIEW", dBIB permite operar referencias individuales de las siguientes formas:

- Ver o listar (a un impresor o a un archivo ASCII) las referencias en formatos de una, dos o tres líneas por referencia y en formato completo con toda la información de la referencia. Cada formato tiene diferentes opciones, variando la disposición de los campos en el formato. Algunos formatos corresponden a standards usados en el NLM.
- Modificar cualquier campo de la referencia. Esto es muy útil, especialmente en el caso de los autores, pues algunos nombres extranjeros a veces están mal escritos. Cuando el usuario modifica el párrafo de palabras clave, puede eliminar o agregar palabras clave a la referencia que se está revisando.
- Marcar la referencia temporalmente: El usuario marca una referencia para agrupar a su voluntad un conjunto de referencias sin tener que utilizar "CRITERIA". En la pantalla la marca se representa con un asterisco (*).
- Anotar la referencia. El usuario puede escribir comentarios acerca de la referencia. Estos comentarios pueden tomarse en cuenta para las

búsquedas del sistema.

Una vez el usuario ha seleccionado un grupo de referencias, ya sea a través del módulo de búsqueda, "CRITERIA", o marcándolas una por una en "VIEW", dBIB permite:

- Cambiar globalmente algunos campos de las referencias, eliminar las referencias de la bibliografía, agregarles palabras clave, etc. La habilidad de agregar palabras clave a un grupo de referencias es muy importante porque así la bibliografía se vuelve dinámica. El usuario tiene la capacidad de establecer y destruir relaciones entre las referencias, liberándose de la rigidez de los "Mesh Headings" del NLM.
- Ordenar las referencias por algún campo como autor, número de identificación de la referencia, nombre de la revista, etc.

8. Mantenimiento de archivos. dBIB ofrece una serie de rutinas que permiten el manejo de las bibliografías, ejecutando procesos como creación y eliminación de bibliografías, copia de respaldo o "backup" y recuperación del respaldo o "restore". También proporciona facilidades para el mantenimiento de los archivos del sistema, como la eliminación física de las referencias borradas por el usuario y la regeneración de los índices físicos y de las relaciones entre los archivos del sistema.

9. Módulos y opciones. En la Figura 12 se ilustra gráficamente todos los módulos y las opciones que dBIB ofrece al usuario.

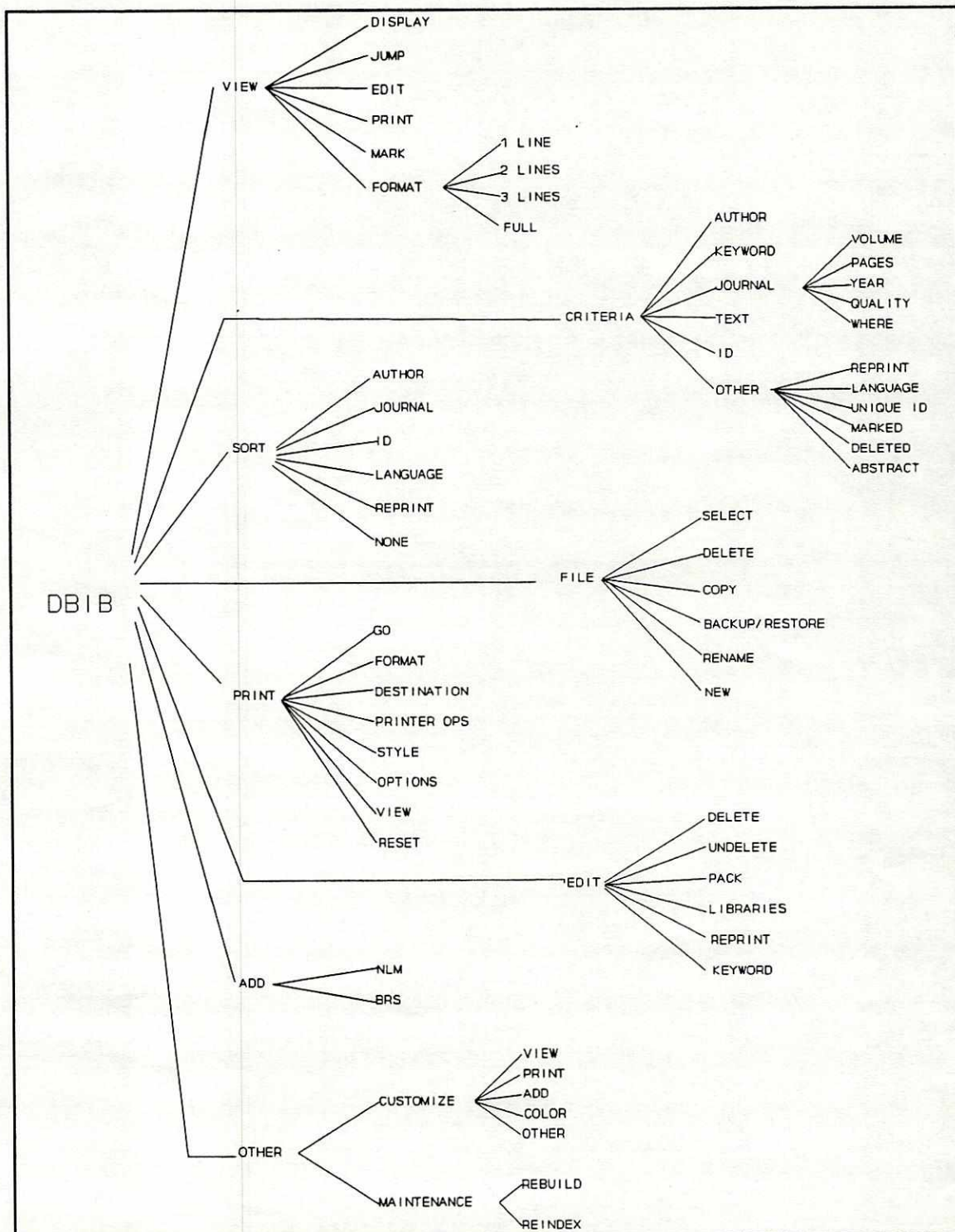


Figura 12. Módulos y submódulos de dBIB.

IV. IMPLEMENTACION DEL SISTEMA

A. Descripción general

dBIB es un sistema computarizado de información que consta de ocho módulos principales. Para su implementación se utilizaron los comandos y rutinas que ofrece CLIPPER, versión 5.01 de Nantucket. CLIPPER es un manejador de bases de datos fundamentado en DBASE¹, orientado al modelo relacional². En su conjunto todos los módulos suman alrededor de 11,000 líneas de código, distribuidos en 50 archivos fuente de programas. El tamaño total de todo el programa presentaba dificultades para su ejecución, pues los computadores personales tienen un máximo de 640 "Kilobytes" memoria convencional. Este problema se solucionó aprovechando una de las ventajas que ofrece CLIPPER: "overlays"³. Utilizando esta técnica el sistema se puede partir en módulos más pequeños llamados "overlays".

En el programa principal, el usuario escoge una de las 8 funciones generales del sistema. En este programa se incluyen las rutinas globales del sistema, tales como rutinas de alerta, definición de formatos de pantalla, mensajes de error, etc.

A continuación se presentan las estructuras de archivos del sistema, se explica la manera en que dBIB maneja algunas funciones de los DRS y se

¹ DBASE es un manejador de bases de datos para computadores personales del tipo IBM. Es un estándar dentro de las oficinas de salud de los Estados Unidos en computadores personales.

² Ver capítulo II.

³ "Overlays": mecanismo que permite compartir la misma área de memoria entre distintas porciones de aplicaciones.

detallan los principales módulos en el sistema.

B. Archivos de dBIB

La Figura 13 muestra los archivos principales que forman una bibliografía en dBIB. Las estructuras de estos archivos se detallan en el Apéndice A.

1. **MAIN: Datos generales de las referencias.** El archivo principal de una bibliografía se llama MAIN. Contiene los datos generales a cada referencia, tales como identificación de la referencia (IDS e IDO), si la referencia proviene de un libro o una revista (SOTYPE), fecha de publicación (DP y MO), volumen y número de páginas (VP), lenguaje original de la referencia (LA), si existe un abstracto (AB), si existe una reimpresión (RP) y la revista en la que aparece (IDR).

La diferencia entre IDS e IDO es que IDS lo asigna dBIB y se utiliza para relacionar internamente el archivo MAIN con otros archivos. En cambio, IDO lo asigna el usuario y él lo utiliza para identificar a su manera las referencias.

IDS e IDR contienen códigos de identificación, de la referencia y de la revista respectivamente. A través de estos códigos se establecen las relaciones entre los archivos. IDS relaciona a MAIN con el archivo de autores AUTHOR y de textos TITLE. IDR por su parte lo relaciona con el archivo de revistas JOURNAL.

Existen otros campos (AUNUM, TINUM, SONUM) en MAIN que también establecen relaciones con los archivos de autores, textos y revistas.

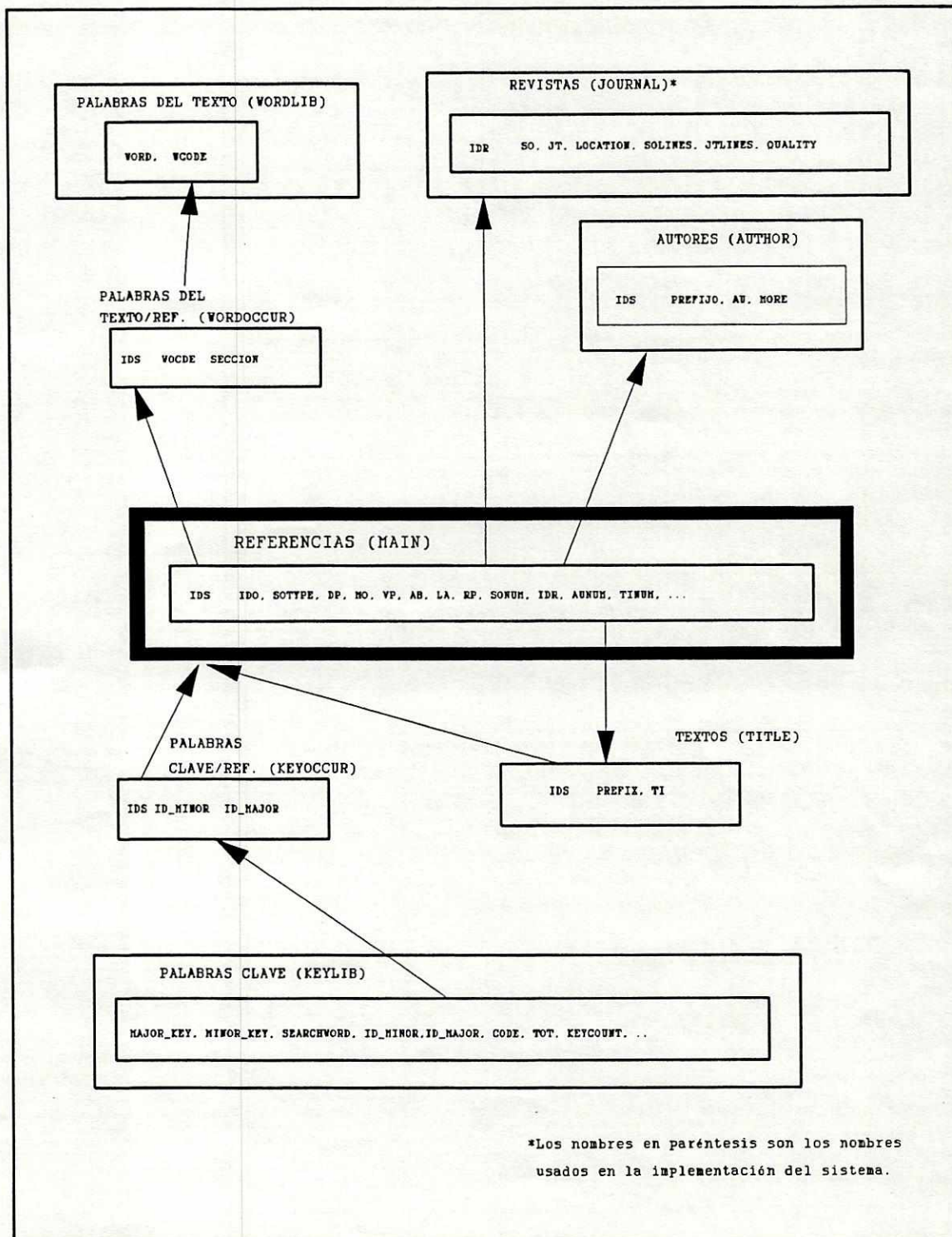


Figura 13. Archivos principales de dBIB.

Estos campos contienen direcciones que apuntan directamente a los registros de los archivos. Por ello permiten obtener la información de estos archivos más rápidamente que por medio de las relaciones que forman IDS e IDR.

El archivo MAIN y los archivos de revistas, autores y textos contienen la información bibliográfica que el usuario frecuentemente revisa a través del módulo "VIEW". De ahí la importancia de presentar esta información lo más rápido posible. Sin embargo, los campos IDS e IDR se usan en los módulos "CRITERIA" y "OTHER/MAINTENANCE" del sistema, según se verá más adelante.

2. **Archivo de textos.** El archivo de textos TITLE contiene los párrafos del abstracto, título original, título en inglés, dirección del lugar donde se escribió el artículo, etc. Estos párrafos se encuentran identificados por el campo PREFIX. El campo TI contiene el texto en formato libre, tal como se encuentra en la referencia original del NLM.

3. **Archivo de autores.** El archivo de autores contiene todos los autores de cada referencia en la bibliografía. El campo PREFIX indica si es el primer autor de la referencia o no. En AU está el nombre del autor. MORE se usa cuando un nombre no cabe en el espacio que AU proporciona. MORE indica que el nombre está incompleto y que continúa en el próximo registro.

El archivo de autores se considera como un diccionario¹, aunque no cumple estrictamente con la característica de no repetir registros (en este caso

¹ Ver capítulo III.

autores) en el archivo. Se investigó la frecuencia de los artículos y autores en las referencias del NLM sobre el tema de leishmaniasis correspondientes a un período de dos años. Esta investigación demostró que en los dos años un autor aparecía en un promedio de 1.8 artículos. Debido a este número tan bajo de repeticiones y al espacio de almacenamiento y tiempo de procesamiento necesarios para mantener un "verdadero" diccionario (archivo adicional que no incluya registros repetidos) de autores, además del archivo AUTHOR, se concluyó que era más conveniente utilizar el archivo AUTHOR como tal y como un diccionario.

4. Archivo de revistas. El archivo de revistas JOURNAL contiene los nombres de las revistas donde se publican las referencias de la bibliografía. Incluye los campos: SO, que contiene el nombre abreviado de la revista; JT, que contiene el nombre completo; LOCATION, que contiene el nombre de la biblioteca donde se encuentra la revista. Cuando el nombre de la revista no cabe en los campos SO y JT, los campos SOLINES y JTLINES indican cuántas líneas o registros más ocupan.

Por último, el campo QUALITY contiene la clasificación subjetiva del usuario referente a la revista: si es mala, regular o buena.

5. Archivos para las palabras clave y palabras del texto. El diccionario de palabras clave KEYLIB contiene cada palabra clave de la bibliografía en los campos MAJOR_KEY y MINOR_KEY, los cuales se explican en otra sección más adelante. En TOT se lleva el número de veces que la palabra clave aparece en la bibliografía. El campo SEARCHWORD

contiene la palabra o palabras de uso común que equivale a la palabra clave definida por MAJOR_KEY y MINOR_KEY. SEARCHWORD se explica en una sección más adelante. El número de palabras en SEARCHWORD se lleva en KEYCOUNT. CODE indica el párrafo de la referencia en el cual es válido encontrar el SEARCHWORD.

Los códigos de identificación de la palabra clave se guardan en ID_MINOR e ID_MAJOR. A través de estos campos se relaciona este archivo con el archivo de palabras clave por referencia KEYOCCUR. En KEYOCCUR se guarda la información de cada vez que aparece la palabra clave en la bibliografía. Se incluye la identificación de la referencia donde aparece o IDS, y la identificación de la palabra clave ID_MINOR e ID_MAJOR.

El diccionario de palabras del texto WORDLIB contiene la palabra en sí WORD y su código de identificación WCODE. A través de este código se relaciona este archivo con el archivo de palabras del texto por referencia WORDOCCUR. En WORDOCCUR se guarda la información de cada vez que aparece la palabra del texto en la bibliografía. Se incluye la identificación de la referencia donde aparece o IDS, la identificación de la palabra WCODE y el campo SECCION, que contiene el párrafo donde aparece la palabra en la referencia.

6. Otros archivos del sistema. Existen en dBIB otros archivos internos necesarios para el funcionamiento del sistema. BIBCAT es un archivo que contiene información sobre las bibliografías. Existe un registro en

BIBCAT por cada bibliografía en el sistema. BIBCAT contiene los nombres físicos y localización de todos los archivos que forman cada bibliografía. Este archivo guarda las opciones que el usuario escoge al momento de crear la bibliografía; por ejemplo, si el usuario desea que en la bibliografía se manejen palabras clave y/o palabras del texto y en qué párrafos de las referencias deben buscarse dichas palabras. El usuario también indica si el archivo puede ser eliminado o si está protegido. Algunos archivos suelen estar protegidos, principalmente los diccionarios, porque varias bibliografías los comparten entre sí. Los archivos pueden ser borrados sólo si son de uso exclusivo de la bibliografía.

El diccionario de palabras sin significado STOPWRDS se usa al agregar referencias a dBIB para evitar que estas palabras sin significado sean agregadas al diccionario de palabras del texto.

El archivo PRINTERS contiene información sobre las impresoras disponibles en el sistema.

dBIB maneja un archivo temporal CRITFILE que contiene los IDS de las referencias seleccionadas en el módulo "CRITERIA". Los valores de las opciones que el usuario elige al usar el sistema también se guardan en archivos. De tal modo, la siguiente vez que use dBIB las opciones en el sistema serán las últimas que él eligió.

C. Descripción de las referencias

Una parte muy importante en los DRS es describir el tema o temas que

trata el documento o referencia. dBIB utiliza dos métodos para describir las referencias bibliográficas: palabras clave y palabras del texto.

1. Palabras clave. Estas palabras las define un experto. En el caso de dBIB, los expertos son el NLM y el científico usuario del sistema. Las referencias bibliográficas del NLM incluyen un párrafo, "Mesh Headings", donde se listan las palabras clave que el NLM eligió para esa referencia.

El usuario puede solicitar a dBIB una lista de todas las palabras clave que aparecen en el párrafo "Mesh Headings" para un grupo de referencias. El NLM publica una serie de tomos de estas palabras clave, donde además de la palabra clave se incluye el SEARCHWORD.

Los SEARCHWORD son las palabras comunes que se encuentran en los párrafos del texto de las referencias como el abstracto, equivalentes a las palabras clave más técnicas. Se usan para asignar palabras clave a las referencias. La diferencia entre SEARCHWORD y palabra clave es que la primera es de uso común en la escritura, mientras la segunda es más técnica y con propósitos de clasificación. Ocasionalmente los SEARCHWORD coinciden con la palabra clave a que equivalen.

Con estos tomos y el listado de dBIB, el usuario puede formar su propio diccionario de palabras clave. Se recomienda que el usuario revise las palabras clave de las referencias que el NLM le envía cada dos o tres meses. Con el tiempo, las palabras claves empiezan a repetirse y la cantidad de nuevas palabras disminuye marcadamente. Es de notar, entonces, que el usuario puede modificar el diccionario de acuerdo a sus necesidades.

La Figura 14 muestra los archivos usados en dBIB para manejar el diccionario de palabras clave.

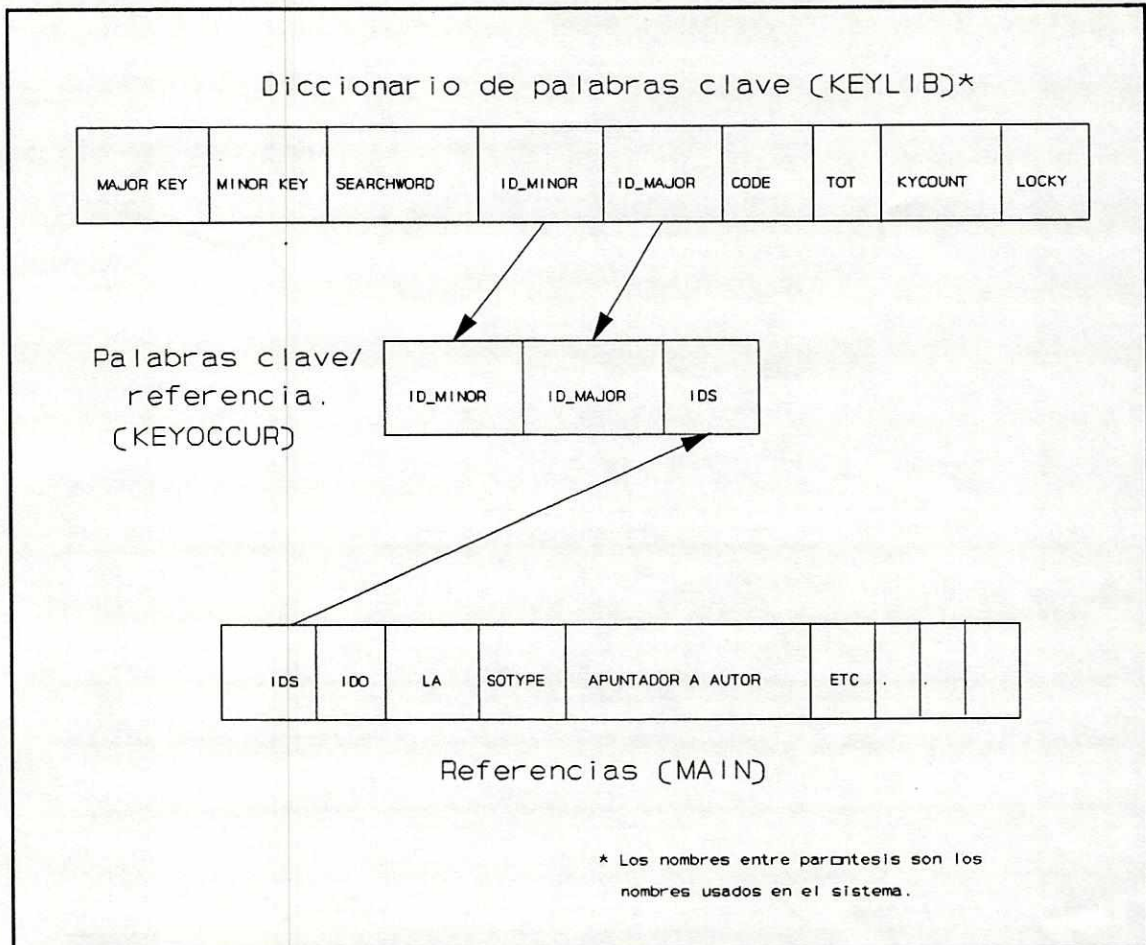


Figura 14. Archivos para palabras claves en dBIB.

El diccionario de palabras clave KEYLIB incluye dos campos MINOR_KEY y MAJOR_KEY, que contienen a la palabra clave. Así cada palabra clave está formada por dos palabras. La palabra en MAJOR_KEY es general, mientras que la palabra en MINOR_KEY es específica.

Según el NLM, esta manera de definir las palabras clave permite precisar mejor el tema de las referencias bibliográficas. En el cuadro de la Figura 15

se ilustra un ejemplo con dos MAJOR_KEY, "TREATMENT" y "DIAGNOSIS". Para "TREATMENT" se establecen los MINOR_KEY: "TREATMENT", "ITRACONAZOLE" y "PLACEBO". Para "DIAGNOSIS" aparecen "CLINICAL" y "PLACEBO". Así, la primera palabra clave es "TREATMENT/TREATMENT", para la segunda es "TREATMENT/ITRACONAZOLE", etc. Una palabra clave es única cuando está definida por ambos MAJOR_KEY y MINOR_KEY. En este ejemplo "PLACEBO" aparece para dos diferentes MAJOR_KEY, definiéndose dos palabras clave.

MAJOR_KEY	MINOR_KEY
TREATMENT	TREATMENT
TREATMENT	ITRACONAZOLE
TREATMENT	PLACEBO
DIAGNOSIS	CLINICAL
DIAGNOSIS	PLACEBO

Figura 15. Palabras clave en dBIB.

El diccionario de palabras clave también incluye al campo SEARCHWORD, que contiene la palabra de uso común que se espera encontrar en el texto de la referencia bibliográfica y que equivale a la palabra clave definida por MAJOR_KEY y MINOR_KEY.

Cuando las referencias son descritas a través de las palabras clave, existe

la posibilidad de que a algunas referencias no les corresponda correctamente ninguna palabra clave. Esto se debe principalmente a que el experto que escoge las palabras clave del sistema realiza un mal trabajo.

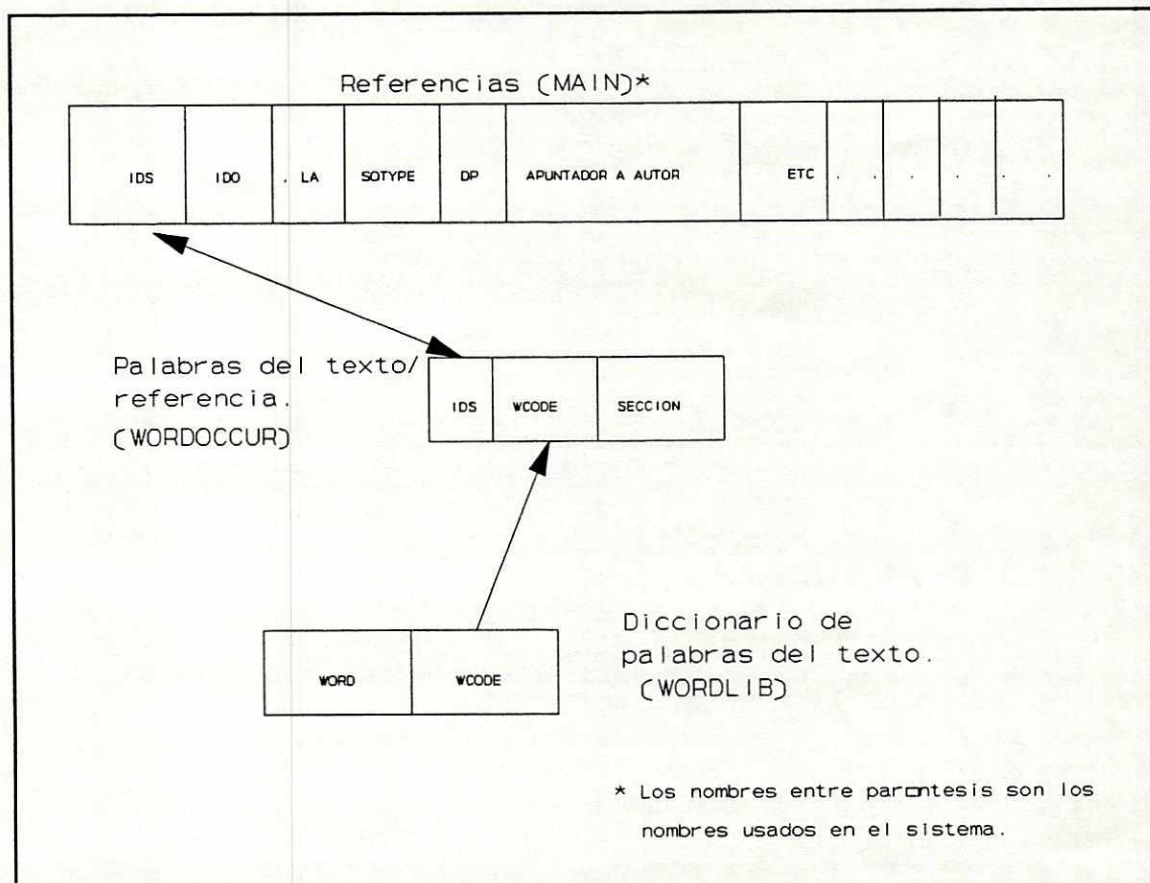


Figura 16. Archivos para palabras del texto en dBIB.

1. **Palabras del texto.** La segunda manera que dBIB ofrece para describir las referencias complementa a la situación anterior, especialmente si se trata de un experto *malo*. dBIB mantiene un registro (ver Figura 16) de todas las palabras con significado que aparecen en el texto de las referencias. Por un lado está el diccionario WORDLIB, que contiene estas

palabras del texto. Por otro lado, en el archivo WORDOCCUR se registra cada vez que aparece una palabra del texto en la referencia. Así cada referencia está asociada a su correspondiente grupo de palabras con significado.

Un inconveniente en la creación del diccionario de palabras de texto es que dBIB incluye todas las que no estén en el diccionario STOPWRDS (palabras sin significado). Esto quiere decir que el programa no puede detectar sinónimos para reducir la redundancia en el diccionario de palabras de texto, ni puede detectar palabras mal escritas. En la práctica se ha observado que aproximadamente el 10% de las palabras que los DRS asignan a los documentos tienen errores de escritura [2].

El sistema STAIRS elimina las variantes gramaticales del diccionario de palabras. Es decir, el diccionario consta de raíces gramaticales en vez de palabras completas. Las palabras que el usuario especifica en el criterio de selección las transforma a su raíz gramatical. El sistema busca la palabra transformada en el diccionario de raíces gramaticales. Si la palabra está mal escrita, o sea que STAIRS no encuentra en el diccionario ninguna raíz gramatical que le corresponda a la palabra, el programa intentará transformarla a otras raíces similares. STAIRS procederá a buscar los documentos que contengan las varias raíces transformadas.

En cambio en dBIB, el control de sinónimos y errores ortográficos lo ejerce el usuario. dBIB provee rutinas rápidas y amigables para que el usuario revise las palabras en el diccionario, escoja cuáles utilizar en el criterio de selección y le permite eliminar palabras del diccionario a su criterio. Para

manejar las palabras afines entre sí (sinónimos), el usuario también puede asignar palabras clave a grupos de referencias y así establecer similitudes entre las referencias. Además, tiene la libertad de agrupar las palabras en el criterio de selección de referencias, formando expresiones lógicas según sus necesidades.

D. Asignación de códigos de identificación

En los archivos KEYOCCUR y WORDOCCUR se guardan las repeticiones de las palabras del texto y las palabras clave para cada referencia. El tamaño de estos archivos es proporcional al número de referencias en la bibliografía. Por ello es necesario que los códigos de identificación en estos archivos sean cortos para que el tamaño de estos archivos sea lo más pequeño posible. Los campos de códigos de identificación de: el documento (IDS), la revista (IDR), las palabras clave (ID_MINOR e ID_MAJOR) y las palabras con significado (WCODE) comparten el mismo método de asignación de códigos.

Todos estos campos son de tipo *character* y varían de dos a cuatro "bytes". Cada "byte" puede tener uno de 222 posibles valores. Estos valores van del ASCII 32 (espacio) al ASCII 253 (carácter gráfico). No se usan los valores ASCII menores de 32 ni mayores de 253 porque CLIPPER reserva algunos de estos valores para su uso interno.

El método de asignación funciona así: Por ejemplo, el código de identificación para las palabras del texto es WCODE y tiene una longitud de tres "bytes". A la primera palabra del texto que se ingrese al diccionario se le

asigna el código {ASCII 32, ASCII 32, ASCII 32}. A la segunda se le asigna {ASCII 32, ASCII 32, ASCII 33}. El tercer "byte" va aumentando en uno el valor ASCII. Los dos primeros "bytes" se quedan con el mismo valor. Al llegar el tercer "byte" al ASCII 253, el siguiente código a asignar es {ASCII 32, ASCII 33, ASCII 32}. El primer "byte" se queda igual, el segundo aumenta en uno el valor ASCII y el tercer vuelve a tener el valor 32. Después de otras 222 palabras, el segundo byte cambia el valor 34. Así, la serie de códigos que pueden ser generados para WCODE es la siguiente:

```

{ASCII 32, ASCII 32, ASCII 32}
{ASCII 32, ASCII 32, ASCII 33}
.
.
{ASCII 32, ASCII 32, ASCII 253}
{ASCII 32, ASCII 33, ASCII 32}
.
.
{ASCII 32, ASCII 33, ASCII 253}
{ASCII 32, ASCII 34, ASCII 32}
.
.
.
{ASCII 32, ASCII 253, ASCII 253}
{ASCII 33, ASCII 32, ASCII 32}
.
.
.
{ASCII 32, ASCII 253, ASCII 253}

```

Como WCODE es de tres "bytes", pueden generarse hasta $222 \cdot 222 \cdot 222$ ó 10,941,048 códigos diferentes. En general, para un campo de N "bytes" se

pueden generar N³ códigos distintos.

Los tamaños de los campos que guardan los nombres de autores (AU), nombres de revistas (SO, JT) y palabras del texto (WORD) se eligieron de tal forma que más del 90% de los posibles valores cupieran en el campo de un sólo registro. Para elegir el tamaño, de estos campos se examinaron dos años de referencias. Se obtuvieron todas las palabras del texto, los nombres de las revistas y los nombres de los autores para este grupo de referencias. Se calcularon sus respectivas frecuencias con base en el número de caracteres que los formaban. De las frecuencias obtenidas se decidió la longitud que debían tener estos campos.

E. Módulos principales de dBIB

1. **Módulo "ADD"**. A través de este módulo se transfiere la información de textos ASCII recibida del NLM y se organiza en el grupo de archivos que forman una bibliografía en dBIB. El módulo se divide en dos partes.

En la primera fase se definen las rutinas generales de pantallas para todos los submódulos de "ADD". En esta fase se interactúa con el usuario para obtener la información necesaria para agregar nuevas referencias a la bibliografía. Se pide al usuario que escoja los archivos del NLM que se van a agregar y que defina el patrón que él desea para el código de identificación de las referencias IDO. Este código consta de diez caracteres, como se muestra en la Figura 17. El IDO se divide en dos partes, separadas por un

punto. La primera está formada por seis caracteres alfanuméricos ("P") y la segunda a la derecha del punto, exclusivamente por tres números ("9"). Para cada archivo del NLM que se va a agregar, el usuario puede decidir cuál es el valor inicial de IDO. Por ejemplo, al agregar el archivo del NLM de nombre LE9001.ASC, el primer valor de IDO puede ser LE9001.001, el segundo LE9001.002 y así sucesivamente. Este ejemplo es un patrón estándar que usan los científicos en MERTU/G. El valor LE9001.001 indica al usuario de MERTU/G que es la primera referencia del archivo de leishmaniasis (indicado por las dos primeras letras LE), del año 1990 (por los dígitos 90) y del mes de enero (por los siguientes dos dígitos 01).

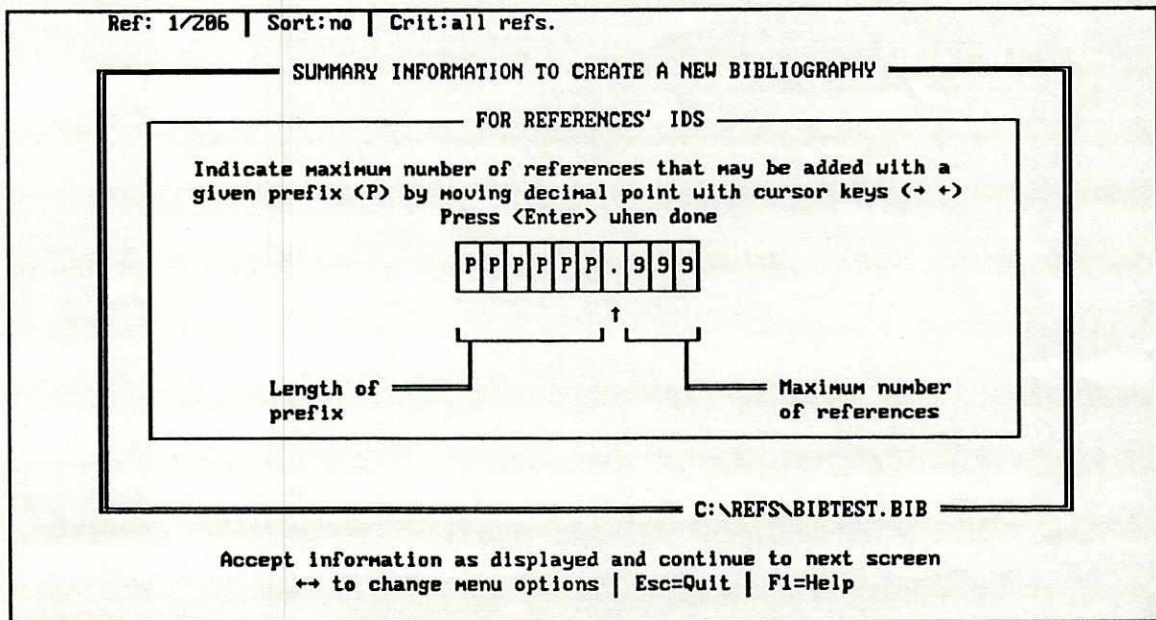


Figura 17. Determinando el patrón para el campo IDO.

En la segunda parte se lee línea por línea cada uno de los archivos del NLM especificados. Se procesa cada palabra de acuerdo al párrafo donde se

encuentra y, según las condiciones en que fue creada la bibliografía, se agrega al archivo correspondiente de la bibliografía. Al crear la bibliografía se especifican ciertas condiciones para el ingreso de las referencias. Estas condiciones indican si dBIB debe formar o no el diccionario de palabras del texto y si debe buscar o no palabras clave en la referencia.

En general, cuando dBIB maneja el diccionario de palabras del texto, el programa lee el texto palabra por palabra. Luego revisa si ya está ingresada al diccionario o no. Si lo está, la descarta y si no, la agrega. El mismo proceso se realiza para el diccionario de palabras clave. En el caso de palabras clave, sin embargo, el proceso es un poco más complejo. Una serie consecutiva de palabras del texto pueden representar una palabra clave. Al no encontrar una palabra individual en el diccionario de palabras clave, dBIB debe revisar si ésta es parte de una secuencia de palabras que representan una palabra clave. Para ello debe obtener las palabras siguientes en la referencia para revisar si la secuencia de palabras está en el diccionario. Si la palabra o palabras corresponde a una palabra clave del diccionario la agrega al archivo KEYOCCUR (apariciones de palabras clave por referencia).

"ADD" utiliza la misma rutina de lectura para los distintos tipos de párrafos que hay en las referencias. Esta rutina obtiene una palabra de la referencia. Dependiendo del párrafo que se trate, esta palabra puede ser el nombre de un autor, de una revista, una palabra clave, etc. La rutina utiliza diferentes grupos de caracteres, para el tipo de párrafo, que delimitan las palabras en el párrafo. Por ejemplo, en el caso de autores, el delimitador de

palabras o nombres es ";". En el párrafo de "Mesh Headings" los delimitadores son "*", "/", " ". Dependiendo del párrafo al que pertenece, la palabra se asigna al archivo correspondiente. Se revisa el caso especial en que la palabra sea más larga que el tamaño del campo donde debe ser almacenada, en cuyo caso el excedente se coloca en un registro adicional.

Por último, "ADD" asigna la demás información bibliográfica a los archivos de dBIB, actualizando los índices y relaciones respectivos.

2. **Módulo "FILE"**. Este módulo se encarga de tareas como creación, copia, eliminación, selección, etc., de las bibliografías. Incluye los siguientes submódulos:

- "NEW". Para crear una bibliografía nueva (y vacía). En "NEW" se deciden las condiciones iniciales de la bibliografía respecto del uso de palabras claves y la creación y uso del diccionario de palabras del texto. Se eligen los nombres físicos de los archivos que forman la bibliografía (autores, títulos, revistas, etc.). Esta información se agrega al archivo BIBCAT.
- "SELECT". Para elegir la bibliografía con la que se quiere trabajar. Al usuario se le muestran las bibliografías existentes en el archivo BIBCAT.
- "DELETE". Elimina una bibliografía. "DELETE" no borra los archivos de los diccionarios que estén compartidos por otra bibliografía.
- "COPY". Copia la bibliografía a una nueva bibliografía. El usuario

elige si quiere copiar los diccionarios o desea compartirlos con la bibliografía nueva.

- "BACKUP/RESTORE". Permite realizar copias de respaldo de la bibliografía a "diskette". Cuando se utiliza la opción "RESTORE", se revisa si en el disco duro ya existe alguna bibliografía con el mismo nombre, para luego renombrarla o reemplazarla.
- "RENAME". Permite cambiar el nombre de la bibliografía.

3. **Módulo "VIEW"**. Es la rutina por la cual el usuario puede ver las referencias en diferentes formatos. Permite modificar la referencia que se está viendo.

En la práctica se ha observado que la mayor parte del tiempo el usuario pasa viendo de una referencia a otra, se detiene a leerla, y vuelve a pasar a otra, se detiene y la examina. Se calcula que en promedio el usuario pasa a ver las referencias *hacia adelante* un 60% más frecuentemente que *hacia atrás*. "VIEW" aprovecha esta situación de la siguiente manera. Mientras el usuario está examinando una referencia no toca el teclado, "VIEW" automáticamente inicia un proceso por medio del cual se leen las referencias en los archivos del disco duro y se pasan a una matriz de referencias en la memoria principal. La Figura 18 ilustra el resultado de este proceso. La pantalla del módulo "VIEW" es como una ventana a esta matriz. Siempre se mantiene un 60% de referencias *hacia adelante* de la pantalla. Este método proporciona una respuesta rápida al usuario al examinar las referencias, pues se evita que el sistema vaya al disco duro a leer las referencias si éstas ya están en la

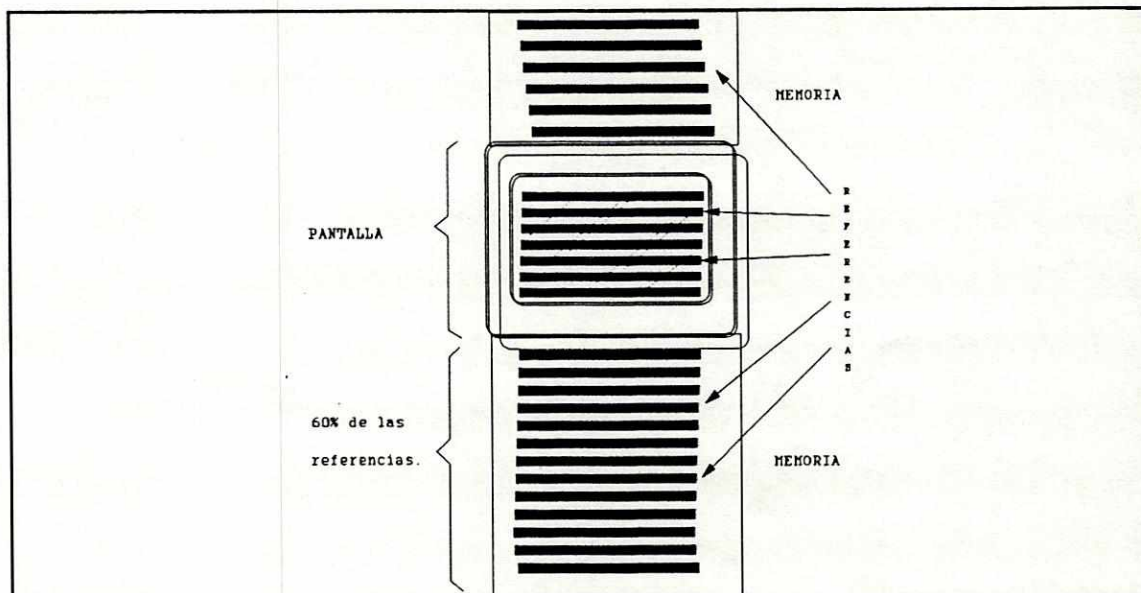


Figura 18. Matriz de referencias en la memoria.

memoria.

En "VIEW" se puede revisar cada referencia con todos sus atributos, hacer cambios en la referencia, incluso eliminarla y hacer anotaciones personales sobre la referencia. Existe la posibilidad de señalar o marcarlas para luego agruparlas en el módulo "CRITERIA".

Los formatos de presentación de las referencias en "VIEW" son de cuatro tipos: de una línea, de dos líneas, de tres líneas por referencia y toda la referencia. Para cada tipo de formato existen diferentes opciones. "VIEW" utiliza la misma rutina para desplegar la información en la pantalla para todos los formatos, menos para el de toda la referencia.

4. Módulo "CRITERIA". Este módulo selecciona referencias de acuerdo a criterios de selección formados por el usuario. Está dividido en dos partes. La primera parte es la interfase con el usuario donde se forma

el criterio de selección. La segunda es la ejecución de dicho criterio y la subsecuente formación del archivo resultado que contiene los códigos de identificación de las referencias encontradas.

La interfase con el usuario es un menú que permite al usuario elegir los atributos de la bibliografía, como autores, nombres de revistas, palabras clave, etc. Además permite elegir los conectores lógicos "AND", "OR" y "NOT" y los paréntesis para formar el criterio de selección. Cuando se eligen atributos como autores, se presentan todos los autores ordenados alfabéticamente en una lista para que el usuario los elija. En este momento el usuario puede elegir, si lo prefiere, a un grupo de autores consecutivo. Esto es muy útil porque a veces los nombres de los autores están mal escritos. Al seleccionar un grupo de variantes del nombre, el usuario se asegura que se incluirán todas las referencias del autor. Una lista similar aparece cuando se eligen palabras clave, palabras del texto, nombres de revistas e identificación de las referencias. En el caso de palabras del texto, si no existe el diccionario correspondiente, no aparecerá la lista ordenada. En su lugar, el programa esperará a que el usuario teclee la palabra que desee buscar. En este caso la búsqueda se realiza secuencialmente línea por línea en el texto de cada referencia.

La segunda parte de este módulo es la búsqueda de las referencias que cumplan con el criterio de selección. "CRITERIA" identifica la primera expresión en el criterio de selección a ser evaluada. Crea un archivo intermedio que contiene los códigos de identificación de las referencias (IDS)

que cumplen con la expresión evaluada. Si la próxima expresión a evaluar está conectada a la primera con un conector lógico "AND", entonces se evalúa la expresión tomando al archivo intermedio como el universo de las referencias. Si la próxima expresión está conectada por un conector "OR", entonces se evalúa y se crea un segundo archivo intermedio que se agrega al primero, eliminándose las referencias repetidas.

Cuando las expresiones están conectadas por un "AND", "CRITERIA" revisa si alguna tiene un índice asociado al campo de la expresión para así evaluar primero esa expresión más rápidamente. La siguiente búsqueda se realizará en un universo de referencias más pequeño.

Una vez realizada la búsqueda, "CRITERIA" permite construir nuevos criterios de selección usando como base el resultado de la búsqueda anterior.

5. Módulo "EDIT". Todos los archivos bibliográficos de dBIB se pueden modificar invocando este módulo. Se puede inclusive escribir una nueva referencia. Se pueden cambiar los diccionarios de la bibliografía, agregar o eliminar elementos del diccionario y agregar nuevas palabras clave. A diferencia de "VIEW", se pueden eliminar todas las referencias que estén seleccionadas en ese momento, o asignarles una palabra clave.

6. Módulo "SORT". dBIB mantiene algunos archivos internos de índices que se pueden usar en cualquier momento. Los índices son sobre los campos: autor, nombre de la revista, identificación del usuario para la referencia, lenguaje en que fue escrito y código de reimpresión. En este módulo se elige qué índice utilizar. En el caso que exista un criterio de

selección activado se crea un índice temporal sobre el archivo resultado de la búsqueda.

7. **Módulo "PRINT"**. Las referencias seleccionadas pueden ser impresas en este módulo. Incluye diferentes opciones para utilizar diferentes tipos de impresoras, listar a pantalla, etc. Se pueden definir nuevos impresores a través de un menú que pide los códigos de control del impresor, el puerto al que está conectado, etc.

8. **Módulo "OTHER"**. Este módulo tiene dos submódulos: "CUSTOMIZE" y "MAINTENANCE". En "CUSTOMIZE" se pueden cambiar las opciones predefinidas para los diferentes módulos al inicio de la sesión de dBIB. Se pueden cambiar los colores de los menús y pantallas del sistema.

En "MAINTENANCE" se ofrecen rutinas que permiten generar de nuevo los índices internos de las bibliografías (el de autores, nombre de revista, etc.). Estas rutinas son importantes en el caso que se corrompan accidentalmente. Revisa que los campos de las direcciones a autores, revistas y al archivo de texto estén correctos. Esto se realiza comparando el código de identificación (IDS) del archivo MAIN con el IDS del archivo de autores y el archivo de texto. En el caso del nombre de revistas se compara el IDR de MAIN con el del archivo de revistas.

V. USO DEL SISTEMA BIBLIOGRAFICO dBIB

En este capítulo se muestra un ejemplo de una sesión en dBIB. El proceso de instalación y creación de una bibliografía se explicó en el capítulo anterior. En este ejemplo se utiliza una bibliografía ya existente y se asume que el sistema está funcionando en el disco duro de un computador personal. Para entrar al sistema, el usuario debe estar en el directorio donde se encuentre dBIB y teclear dBIB. Al final de este capítulo se encuentran las figuras a las que se hace mención en los párrafos siguientes.

La Figura 19 muestra el menú principal de dBIB y en él aparece seleccionada la opción "VIEW". El nombre de la bibliografía aparece sobre la línea inferior del marco del menú y corresponde a "C:\REFS\BIBTEST.BIB". Al seleccionar "VIEW" llegamos a la pantalla de la Figura 20 donde aparece elegida la segunda referencia, con autor Albanese G. Aquí se muestra el código de identificación de la referencia asignado por el usuario, todos los autores de la referencia, el título de la referencia, el nombre de la revista donde se publicó, volumen, páginas, fecha de publicación y el abstracto de la referencia. La línea de status, en la parte superior de la pantalla, indica que las 236 referencias de la bibliografía están seleccionadas y que la que se encuentra señalada es la número dos. La Figura 20 muestra el cambio del formato de las referencias. Se elige en este caso una línea por referencia. El cambio se muestra en la Figura 21.

El menú de "CRITERIA" se muestra en la Figura 22. Al elegir la opción

de palabras clave, aparece la Figura 23. En ella se elige la palabra clave "TREATMENT" y el resultado se muestra en la Figura 24. La línea de status indica que 49 referencias cumplen con que la palabra clave sea "TREATMENT". En la figura 25 se muestra la pantalla que aparece al entrar a "CRITERIA" con un criterio de selección activado. La Figura 26 muestra el menú de conectores lógicos. En la Figura 27 se relaciona el criterio anterior, por medio de un "AND", con el criterio de que el autor sea igual a "NAVIN". El resultado de este nuevo criterio de selección son dos referencias (Figuras 28 y 29). En la Figura 30 se termina la sesión con dBIB.

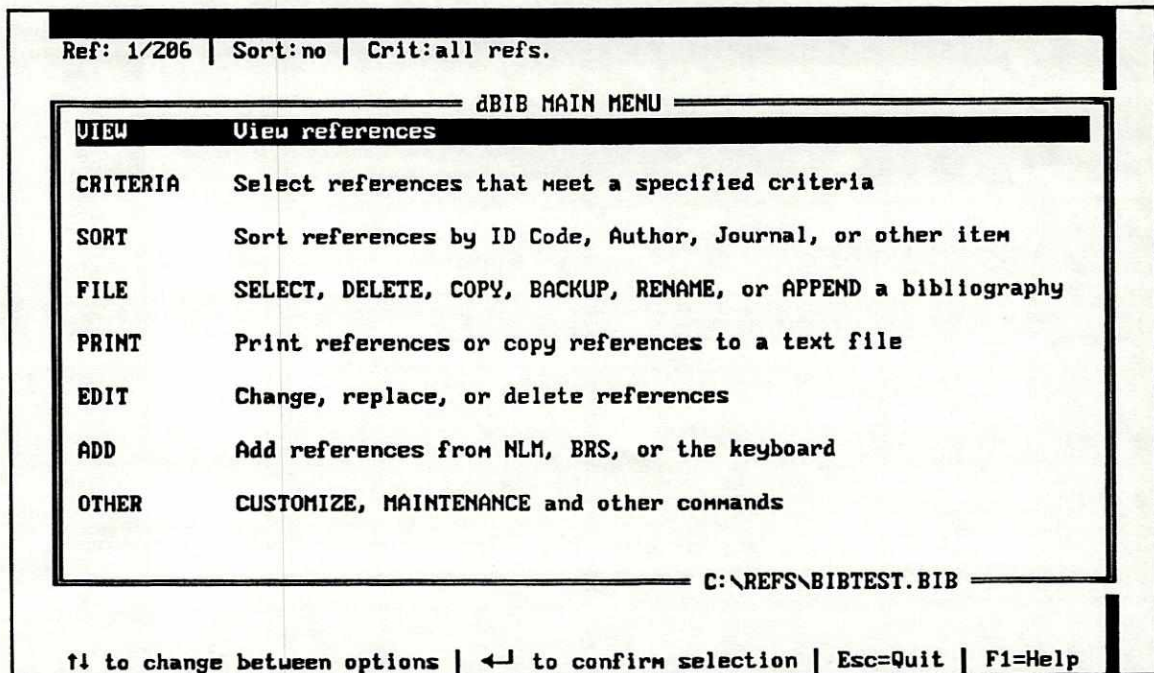


Figura 19. Menú principal.

Ref: 2/206 | Sort:no | Crit:all refs.
 DISPLAY J

Mesa-Ualle	LE9801.002 Albanese G, Giorgetti P, Santagostino L, Crippa D and Sala G. Cutaneous leishmaniasis. Treatment with itraconazole. Arch Dermatol 1989 125(11):1540-2
Albanese G	Abstract
Pratlong F	The chemotherapy of leishmaniasis is still far from satisfactory. Itraconazole, one of the most recent azole antifungal agents, appears to be well tolerated in man. Two patients with cutaneous leishmaniasis (<i>Leishmania tropica</i>), contracted in Saudi Arabia, were treated orally for 2 months with itraconazole (100 mg/d). Both recovered without side effects or abnormalities in their main biologic parameters.
Dourado MI	
Gomes A de	
Cairns BR	

Displays the reference's abstract

Figura 20. Referencias en "VIEW".

Ref: 2/206 | Sort:no | Crit:all refs.

Option	Description
A	ID, Journal, Year, Volume, Pages
B	ID, Title
C	ID, Data
D	ID, Reprint Location
E	First Author, Journal, Year, Volume, Pages
F	First Author, Title
G	First Author, Data
H	First Author, Reprint Location
I	ID

Examples of Format option A

LE8802.002	Ann Trop Med Parasitol	1987 81(2):105-16
LE8804.008	Trop Med Parasitol	1987 38(3):215-21
LE8809.006	Trans R Soc Trop Med Hyg	1986 80(6):859-76

↑↓=move between Format options | ← to confirm | ESC=Quit
 For format options that display 1 line of a reference

Figura 21. Formatos de referencias.

ID	SOURCE	YEAR	VOLUME: PAGES
LE9001.002	Arch Dermatol	1989	125(11):1540-2
LE9001.003	Ann Parasitol Hum Comp	1989	64(4):312-4
LE9001.004	Rev Saude Publica	1989	23(1):2-8
LE9001.005	Rev Inst Med Trop Sao Paulo	1989	31(1):32-9
LE9001.006	Proc Natl Acad Sci U S A	1989	86(20):7682-6
LE9001.007	Parazitologia	1989	23(4):334-43
LE9001.008	Med Parazitol (Mosk)	1989	(4):79-86
LE9001.009	Med Parazitol (Mosk)	1989	(4):33-6
LE9001.010	Med Parazitol (Mosk)	1989	(4):28-33
LE9001.011	Lancet	1989	2(8673):1222-3
LE9001.012	Jpn J Exp Med	1989	59(3):103-8
LE9001.013	J Protozool	1989	36(5):510-8
LE9001.014	J Protozool	1989	36(5):502-10
LE9001.015	J Protozool	1989	36(5):498-501
LE9001.016	J Infect Dis	1989	160(5):904-5
LE9001.017	J Immunol	1989	143(9):3020-4

C:\REFS\BIBTEST.BIB

↑↓=move between Format options | ← to confirm | ESC=Quit
Displays the reference's citation

Figura 22. Referencias en formato de una línea.

TO SELECT ALL REFERENCES	
AUTHOR	Written by a specific author
KEYWORD	Flagged with a specific key word
JOURNAL	From a specific journal, volume, or year of publication
TEXT	Containing a specific word in the title, abstract, etc
ID	By your DBIB identification code
OTHER	By reprint, language, UI, being MARKed, record number, and others
RELATE	Indicate relation between criteria (eg, AND, OR, NOT)
EXECUTE	Carry out selection of references that fulfill your criteria
CRITERIA SELECTED: _____	

↑↓ to change between options | ← to confirm selection | Esc = Quit | Del = Erase

Figura 23. Menú de "CRITERIA".

KEYWORDS	
Major Key	Minor Key
<p>TREATMENT</p> <p>VECTOR</p> <p>VETERINARY DISEASES</p>	
Major Key Speed Search: TR	
↑↓ to change between keywords ← to confirm selection Esc=Quit F1=Help	

Figura 24. Palabras clave.

ID	SOURCE	YEAR	VOLUME: PAGES
LE9001.001	Arzneimittelforschung	1989	39(8):838-42
LE9001.002	Arch Dermatol	1989	125(11):1540-2
LE9001.008	Med Parazitol (Mosk)	1989	(4):79-86
LE9001.012	Jpn J Exp Med	1989	59(3):103-8
LE9001.016	J Infect Dis	1989	160(5):904-5
LE9001.017	J Immunol	1989	143(9):3020-4
LE9001.021	J Assoc Physicians India	1989	37(2):154-6
LE9001.025	Int J Immunopharmacol	1989	11(6):637-45
LE9001.030	C R Acad Sci [IIII]	1989	308(18):485-8
LE9002.001	Rev Inst Med Trop Sao Paulo	1989	83(1):19-24
LE9002.003	Parazitologia	1988	39(5-6):473-4
LE9002.009	J Immunol	1989	143(12):4244-9
LE9002.016	J Immunol	1989	119(38):1309-13
LE9002.024	Int J Parasitol	1989	28(9):612-4
LE9002.027	Indian J Med Res	1989	114(48):1876-81
LE9003.001	J Immunol	1989	69(3):223-32

C:\REFS\BIBTEST.BIB

↔=change menu option. | ↑↓=move between Move options | *=marked ref | ESC=Quit
 Displays the reference's citation

Figura 25. Referencias con la palabra clave "TREATMENT".

TO SELECT ALL REFERENCES	
AUTHOR	Written by a specific author
KEYWORD	Flagged with a specific key word
JOURNAL	From There is an active criteria... you may: ion
TEXT	Code ADD To add items to the current criteria tc
ID	By EDIT To edit the current criteria string
OTHER	By CLEAR To erase the current criteria r, and others
RELATE	In)
EXECUTE	Carry out selection of references that fulfill your criteria
CRITERIA SELECTED: _____	
KW = TREATMENT	
↑↓ to change between options ← to confirm selection Esc = Quit Del = Erase	

Figura 26. Criterio de selección activado.

RELATE	
AND	Restricts search to references that meet both criteria
OR	Includes references that meet either criteria
(Use open and close parenthesis to group criteria
)	eg Author = Neva AND (Keyword = TREATMENT OR Keyword = DIAGNOSIS)
NOT	Use NOT to negate the next search criterion
CS	
↑↓ to change between options ← to confirm selection Esc=Quit F1=Help	

Figura 27. Menú de conectores lógicos.

AUTHOR

ALL AUTHORS

Navin TR

Neal RA

Neva FA

Ngige S

Nieves E

Nishioka S de A

Nunes MP

Nunes UL

Speed Search: **NAUI**

→ From Author: **Navin TR** To Author: **Navin TR**

Notes: 1. Capitalization is disregarded (eg, LAINSON R = lainson r)
 2. Use → ← to switch between FROM Author and TO Author

↑↓ to change between authors | ← to confirm selection | Esc=Quit | F1=Help

Figura 28. Pantalla de autores en "CRITERIA".

ID	SOURCE	YEAR	VOLUME: PAGES
LE9004.044	Immunol Today	1990	42(1):43-50
LE9004.045	Immunol Today	1990	42(1):36-42

C:\REFS\BIBTEST.BIB

←+=change menu option. | ↑↓=move between Move options | *=marked ref | ESC=Quit
 Displays the reference's citation

Figura 29. Pantalla de "VIEW" en formato de una línea.

Ref: 1/2 | Sort: no | Crit: CS AND Author = NAVIN TR
 DISPLAY J

ID	
LE9004.044	Navin TR, Arana BA, Arana FE, de Merida AM, Castillo AL and Pozuelos JL.
LE9004.045	Placebo-controlled clinical trial of meglumine antimonate (glucantime) vs. localized controlled heat in the treatment of cutaneous leishmaniasis in Guatemala. Immunol Today 1990 42(1):43-50

Abstract

Sixty-six Guatemalans with parasitologically proven cutaneous leishmaniasis were randomly and equally divided into 3 treatment groups: those receiving meglumine antimonate (Glucantime), 850 mg antimony/day im for 15 days; those receiving localized controlled heat from a radio-frequency generator, 50 degrees C for 30 sec, 3 treatments at 7 day intervals; and those receiving treatment with a placebo. Of 53 isolates identified, 40 were Leishmania braziliensis braziliensis and 13 were L. mexicana mexicana. Thirteen weeks after beginning treatment, the number of patients from each group with completely healed

Displays the reference's abstract

Figura 30. Abstracto de una referencia.

Ref: 1/2 | Sort: no | Crit: CS AND Author = NAVIN TR

===== dBIB MAIN MENU =====

VIEW	View references		
CRITERIA	Select references that meet a specified criteria		
SORT	Sort refere		l, or other item
FILE	SELECT, DEL	Exit dbib ?	NO PPEND a bibliography
PRINT	Print refer		YES ext file
EDIT	Change, rep		
ADD	Add references from NLN, BRS, or the keyboard		
OTHER	CUSTOMIZE, MAINTENANCE and other commands		

C:\REFS\BIBTEST.BIB

↑↓ to change between options | ← to confirm selection | Esc=Quit | F1=Help

Figura 31. Final de sesión en dBIB.

VI. CONCLUSIONES Y RECOMENDACIONES

Los sistemas de recuperación de información textual o de documentos, DRS, se pueden implementar como un sistema de información dentro del contexto de sistemas manejadores de bases de datos en computadores personales compatibles con IBM. El manejador de bases de datos CLIPPER, orientado al modelo relacional, es una plataforma adecuada para el desarrollo y manejo de sistemas bibliográficos de información, específicamente de dBIB. En el desarrollo de este sistema se aprovecharon las investigaciones existentes sobre métodos de acceso a textos en DRS. El sistema de prototipo para la creación de pantallas y menús satisface al máximo al usuario aunque prolonga el desarrollo del sistema.

A diferencia de los DRS tradicionales, dBIB ofrece una interfase ágil y amigable. Permite que el usuario intervenga en la determinación de sinónimos y en la definición de palabras clave nuevas. De este modo el usuario experto tiene más posibilidades de encontrar las referencias que él desee.

dBIB presenta ventajas sobre otros sistemas similares instalados en computadores más grandes. Se comparó la eficacia de la búsqueda de las referencias correctas del sistema MEDLINE del NLM. Un sistema como dBIB, cuya interfase con el usuario permite que éste interactúe más con el sistema, demostró ser más eficiente que MEDLINE si el usuario conoce el tema que busca.

El sistema provee una manera para uniformar y comunicar referencias entre

usuarios de dBIB. Se pueden compartir los diccionarios de palabras clave personales, transmitiendo a otros no sólo las referencias sino que las relaciones y anotaciones creadas por otro usuario. El sistema dBIB representa una herramienta muy útil para los científicos de MERTU/G y mejora en grandes proporciones el manejo manual de las referencias bibliográficas.

El sistema fue desarrollado alrededor de una estructura predefinida y no permite hacer cambios de tamaños de campos y no existen rutinas que permitan definirlos. Sin embargo las rutinas en dBIB no dependen del tamaño de los campos o de la estructura de los archivos, sino del nombre de los campos. Se propone usar como base los algoritmos de los módulos de dBIB y agregar rutinas para la definición de archivos y la creación de índices y relaciones entre archivos. Lo anterior convertiría al sistema en un generador de sistemas bibliográficos y podría ser usado para cualquier tipo de biblioteca.

Los mensajes y comentarios del sistema pueden traducirse a otro lenguaje para que los usuarios que no sepan inglés puedan utilizar el sistema.

El sistema dBIB en la actualidad está siendo usado por los científicos en MERTU/G y principa a utilizarse por otros científicos en otras estaciones de campo.

VII. BIBLIOGRAFIA

1. AHO, Alfred. V., John. E. HOPCROFT, Jeffrey. D. Ullman. The Design and Analysis of Computer Algorithms. Addison-Wesley, Reading, Mass., 1983.
2. BOURNE, C. P. "Frequency and Impact of Spelling Errors in Bibliographic Databases". Information Process Management. Vol 13, No. 1, 1977. pp. 1-12.
3. CLIPPER Reference's Manual. Nantucket Corporation. 1990
4. CODASYL Report. Data Base Task Group of CODASYL Programming Language Committee. Report. 1971.
5. CRAWFORD, R. G., I. A. MACLEOD. "A Relational Approach to Modular Information Retrieval Systems Design". Proceedings of the 41st Conference of the American Society for Information Science 15. 1980. pp 83-85.
6. DATTOLA, R. M. "Flexible Information Retrieval System for Text". Journal of the American Society for Information Science 30. 1979. pp 29-35.
7. FALOUTSOS, Christos. "Access Methods for Text". ACM Computer Surveys, Vol. 17, No. 1, marzo 1985. pp 49-74.
8. KNUTH, D. E. The Art of Computer Programming, Volume 3: Sorting and Searching. Addison-Wesley, Reading, Mass, 1973.

9. MACLEOD, I. A. "A Data Base Management System for Document Retrieval Applications". Information Systems 6. 1983. pp 131-137.
10. MCCARN, D. B., J. P. LEITER. "On-line Services in Medicine and Beyond". Science 181. 1975. pp 318-324.
11. SALTON, G., M. J. MCGILL. Introduction to Modern Information Retrieval. McGraw-Hill Book Company, New York, 1985.
12. SCHECK, H. J., P. PISTOR. "Data Structures for an Integrated Data Base Management and Information Retrieval System". Proceedings of the Eighth International Conference on Very Large Data Bases. 1984. pp 197-207

APENDICE A

ESTRUCTURAS DE ARCHIVOS EN dBIB

Archivo Main: datos generales de cada referencia.

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
IDS	C,4	Identificación del sistema para el documento.
IDO	C,10	Identificación del usuario para el documento.
SOTYPE	C,1	Contiene valores que especifican si la referencia se publicó en un libro o en una revista.
DP	C,4	Fecha de impresión.
MO	C,7	Mes de impresión.
VP	C,16	Volumen y páginas.
AB	L,1	Tiene abstracto.
LA	C,3	Lenguaje original.
RP	C,1	Tiene reimpresión.
SONUM	N,5	Dirección al archivo JOURNAL.
IDR	C,4	Código de la revista.
AUNUM	N,7	Dirección al archivo de AUTHOR.
TINUM	N,7	Dirección al archivo de TITLE.
MARK	C,1	Campo para marcar la referencia a petición de usuario.
ID_MINOR	C,2	Dirección al archivo KEYLIB, indicando la localización del artículo.

Archivo Author

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
IDS	C,4	Identificación de la referencia.
PREFIX	C,1	Indica si es el primer autor.
AU	C,17	Nombre del Autor.

MORE	C,1	Indica si el nombre continúa en otro registro.
------	-----	--

Archivo Title: párrafos de texto en formato libre de las referencias.

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
IDS	C,4	Identificación de la referencia.
PREFIX	C,1	Prefijo que indica el párrafo al que pertenece el texto en TI.
TI	C,69	Texto en formato libre.

Archivo Journal: revistas de la bibliografía.

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
IDR	C,4	Identificación de la revista.
SO	C,30	Nombre de la revista, versión abreviada.
JT	C,60	Nombre de la revista, versión completa.
LOCATION	C,2	Biblioteca o lugar donde se encuentra la revista.
SOLINES	N,1	No. de líneas que ocupa el nombre abreviado de la revista.
JTLINES	N,1	No. de líneas que ocupa el nombre completo.
QUALITY	C,1	Calidad de la revista.

Archivo Keylib: diccionario de palabras clave.

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
MAJOR_KEY	C,30	Palabra clave general.
MINOR_KEY	C,30	Palabra clave específica.
SEARCHWORD	C,40	Palabra que puede aparecer en el texto.
ID_MINOR	C,2	Código de indentificación para la palabra clave específica.
ID_MAJOR	C,1	Código de indentificación para la palabra clave general.
CODE	C,1	Párrafo donde puede aparecer la palabra SEARCHWORD.
TOT	C,3	Veces que aparece la palabra clave.
KEYCOUNT	N,1	Número de palabras en SEARCHWORD.

LOCKEY	C,1	Indica si la palabra clave es el nombre de una localización.
--------	-----	--

Archivo Keyoccur: palabras clave por cada referencia.

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
IDS	C,4	Identificación de la referencia.
ID_MINOR	C,2	Código de la palabra clave específica.
ID_MAJOR	C,1	Código de la palabra clave general.

Archivo Wordlib: diccionario de palabras del texto.

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
WORD	C,15	Contiene la palabra.
WCODE	C,3	Código de la palabra.

Archivo Wordoccur: palabras del texto por cada referencia.

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
IDS	C,4	Identificación de la referencia.
WCODE	C,3	Código de la palabra.
SECCION	C,1	Párrafo donde se encontró la palabra.

Archivo Notes: comentarios del usuario sobre las referencias.

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
IDS	C,4	Identificación de la referencia.
NOTE	C,69	Texto de la nota.

Archivo Stopwrds: palabras sin significado.

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
STOPWORD	C,15	Palabra sin significado, como pronombres y artículos.

Archivo Printers: información de impresoras.

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
NAMEPRN	C,25	Nombre del impresor.
NORMAL	C,30	Secuencia de "caracteres" que provocan que el impresor use el modo normal al imprimir.
BOLD	C,30	Secuencia de "caracteres" que provocan que el impresor use el modo realzado al

UNDERLINE	C,30	imprimir. Secuencia de "caracteres" que provocan que el impresor use el modo subrayado al imprimir.
NOUNDERLIN	C,30	Secuencia de "caracteres" que evitan que el impresor use el modo subrayado al imprimir.
COMPRESS	C,30	Secuencia de "caracteres" que provocan que el impresor use el modo comprimido al imprimir.
ELITE	C,30	Secuencia de "caracteres" que provocan que el impresor use el modo elite al imprimir.
DRAFT	C,30	Secuencia de "caracteres" que provocan que el impresor use el modo draft al imprimir.
NLQ	C,30	Secuencia de "caracteres" que provocan que el impresor use el modo NLQ (Near Letter Quality) al imprimir.
LQ	C,30	Secuencia de "caracteres" que provocan que el impresor use el modo LQ (Letter Quality) al imprimir.

Archivo Bibcat: información sobre las bibliografías.

<u>Campos</u>	<u>Tipo y tamaño</u>	<u>Descripción</u>
MAINPATH	C,40	Directorio donde se encuentra el archivo MAIN.
MAIN	C,8	Nombre del archivo MAIN.
MAIN_EXT	C,3	Extensión del archivo MAIN.
TIPATH	C,40	Directorio donde se encuentra el archivo TITLE .
TITLE	C,8	Nombre del archivo TITLE.
TITLE_EXT	C,3	Extensión del archivo TITLE.
AUPATH	C,40	Directorio donde se encuentra el archivo AUTHOR.
AUTHOR	C,8	Nombre del archivo AUTHOR.
AUTHOR_EXT	C,3	Extensión del archivo AUTHOR.

JOPATH	C,40	Directorio donde se encuentra el archivo JOURNAL.
JOURNAL	C,8	Nombre del archivo JOURNAL.
JOUR_EXT	C,3	Extensión del archivo JOURNAL.
KEYLPATH	C,40	Directorio donde se encuentra el archivo KEYLIB.
KEYLIB	C,8	Nombre del archivo KEYLIB.
KEYLIB_EXT	C,3	Extensión del archivo KEYLIB.
KEYOCCPATH	C,40	Directorio donde se encuentra el archivo KEYOCCUR.
KEYOCCUR	C,8	Nombre del archivo KEYOCCUR.
KEYOCC_EXT	C,8	Extensión del archivo KEYOCCUR.
WORDLPATH	C,40	Directorio donde se encuentra el archivo WORDLIB.
WORDLIB	C,8	Nombre del archivo WORDLIB.
WORDL_EXT	C,3	Extensión del archivo WORDLIB.
WORDOCPATH	C,40	Directorio donde se encuentra el archivo WORDOCCUR.
WORDOCCUR	C,8	Nombre del archivo WORDOCCUR.
WORDOC_EXT	C,3	Extensión del archivo WORDOCCUR.
NOTESPATH	C,40	Directorio donde se encuentra el archivo NOTES.
NOTES	C,8	Nombre del archivo NOTES.
NOTES_EXT	C,3	Extensión del archivo NOTES.
DATE_INI	DATE,8	Fecha de creación de la bibliografía.
DATE_MOD	DATE,8	Fecha de última modificación.
RECNUM	N,7	Número de referencias en la bibliografía.
PROTECTMAP	C,12	Cuales archivos están protegidos, compartidos con otras bibliografías.
ID_DECIMAL	N,1	Posición del punto en el campo ido.
TI_KW	L,1	Indica si debe buscar palabras clave en el Título.
AB_KW	L,1	Indica si debe buscar palabras clave en el Abstracto.
MH_KW	L,1	Indica si debe buscar palabras clave en el párrafo de "Mesh Headings".
OT_KW	L,1	Indica si debe buscar palabras clave en el

TI_WD	L,1	párrafo de otros. Indica si debe buscar palabras del texto en el Título.
AB_WD	L,1	Indica si debe buscar palabras del texto en el Abstracto.
MH_WD	L,1	Indica si debe buscar palabras del texto en "Mesh Headings".
OT_WD	L,1	Indica si debe buscar palabras del texto en el párrafo de otros.

APENDICE B

MANEJO DE ERRORES INESPERADOS EN dBIB

Aunque las rutinas de captura de errores pueden controlar la mayoría de errores, hay tres tipos de errores incontrolables que pueden ocurrir en dBIB:

1. Errores cuando el impresor no está listo
2. Errores de archivos, tales como: abrir o cerrar archivos inexistentes, errores en archivos de la base de datos al agregar o reemplazar campos que no existen.
3. Errores de bajo nivel del sistema operativo DOS, por ejemplo: disco lleno, memoria principal llena, etc.

Lo siguiente es una descripción de cómo dBIB maneja cada uno de estos errores.

ERRORES CUANDO EL IMPRESOR NO ESTA LISTO.

Ocurren cuando la comunicación con el impresor se interrumpe durante una orden de impresión. Ejemplos típicos son: no hay papel, se apaga el impresor, etc. Cuando esto ocurre, dBIB se detiene, despliega un mensaje de alerta en la pantalla. Permite al usuario arreglar el impresor, continuar imprimiendo o abandonar la impresión.

Notas de ayuda para la programación en CLIPPER.

1. Antes de cualquier operación de impresión, se debe chequear el status del impresor con la función ISPRINTER(). Esta rutina es más rápida que la rutina de recuperación ERRORSYS. ERRORSYS sólo debe invocarse cuando el impresor presenta un problema durante la impresión.
2. Para cada operación de impresión se deben encerrar las instrucciones entre un BEGIN y END de tal forma que si el impresor presenta un problema durante la impresión, el control del programa puede pasar a la rutina de recuperación ERRORSYS.
3. Hay una espera de más de un minuto entre la interrupción de la impresión y la aparición del mensaje de alerta. Este tiempo de retraso lo determina DOS y no puede ser alterado desde CLIPPER.

ERRORES DE ARCHIVOS.

Si inesperadamente algún archivo de dBIB es eliminado, se producirá un error al tratar de ser usado en dBIB. La rutina automática de error ERRORSYS puede manejar esta clase de errores. Muestra un mensaje explicando el error y regresa el control al programa para que siga funcionando. Esta rutina es de mucha ayuda durante el desarrollo de los programas, pues sirve también como una herramienta para determinar errores de programación.

ERRORES DE BAJO NIVEL DEL DOS

Ciertos errores no pueden ser interceptados por las instrucciones del programa del sistema ni por la rutina ERRORSYS. Estos errores incluyen: disco lleno y memoria principal llena. En dBIB se decidió calcular la memoria máxima que el programa necesita para evitar rebasar los 640 KB que tienen la mayoría de computadoras personales y evitar errores fuera de control. En cuanto al disco lleno, cada operación de escritura a disco va precedida de una revisión del espacio disponible en disco.