

// **LAS BASES DE DATOS: UNA HERRAMIENTA MODERNA QUE CONTRIBUIRA
A LA OPORTUNA DIFUSION DE INFORMACION
CIENTIFICA Y TECNICA**

Mario Gutiérrez Jiménez

Trabajo presentado en el Curso de Escritura Técnica y Científica, Sala de Ex-Rectores, Biblioteca Central, Universidad Nacional, Heredia, Costa Rica. 20 al 22 de noviembre de 1985.

La preparación, reproducción y distribución de este documento, es patrocinado por el Programa Suizo de Cooperación para el Desarrollo, DDA, por medio de INFORAT: Información y Documentación Forestal para América Tropical.

**CENTRO AGRONOMICO TROPICAL DE INVESTIGACION Y ENSEÑANZA, CATIE
Departamento de Recursos Naturales Renovables
Turrialba, Costa Rica, 1985**

GUTIERREZ, M. Las bases de datos: una herramienta moderna que contribuirá a la oportuna difusión de información científica y técnica. Turrialba, Costa Rica, CATIE, 1985. 22 p. 11 refs. (mimeogr).

COMPENDIO

Se definen algunos términos usados en informática como: dato, información, procesamiento de datos y base de datos. La información como elemento esencial en la toma de decisiones. Evolución de las metodologías en el manejo de datos. Objetivos, productos y propósitos a alcanzar con las bases de datos. El lenguaje natural y el lenguaje documental. El *thesaurus* como elemento unificador de la terminología utilizada en las bases de datos. Las palabras clave y los descriptores. El ingreso de datos, los campos de interés y el sistema de manejo de una base de datos. Las pruebas de validez, la seguridad y el respaldo que se debe dar a la información ingresada. Las redes computarizadas de bibliotecas. La situación en América Latina con respecto a las bases de datos.

ABSTRACT

Several terms used in information science are defined: data, information, data processing and database. The importance of relevant information in the decision making process; evolution in data management methodologies. Objectives, products and users of databases. Difference between natural language and documental language. The thesaurus as a unifying element of terminology utilized in databases. Key words and descriptors. How to access data; fields of interest and database management system (DBMS). Validity of databases contents and retrieval performance. Backup and security of databases. Computerized libraries networks. The data base situation in Latina America.

INDICE

Página

Compendio	I
Abstract	I
Definición de algunos términos usados en Informática	1
Retrospectiva histórica del procesamiento de información	2
Las bases de datos como herramienta de difusión de nuevos conocimientos	3
a. Definición	3
b. Objetivos	3
c. Productos	4
d. Usuarios	4
Propósitos a alcanzar	4
Necesidades de usar un lenguaje documental internacionalmente aceptado	5
El ingreso de datos	6
a. El sistema de manejo de una base de datos	6
b. La estructura de la base de datos	7
c. Campos de interés	7
d. Pruebas de validez de la base de datos	8
e. Limitaciones de las bases de datos	9
f. Seguridad y respaldo	9
Las redes computarizadas de bibliotecas	9
La situación en América Latina en relación con las bases de datos	10
Agradecimientos	11
Literatura consultada	11

LAS BASES DE DATOS: UNA HERRAMIENTA MODERNA
QUE CONTRIBUYA A LA OPORTUNA DIFUSION
DE INFORMACION CIENTIFICA Y TECNICA *

Mario Gutiérrez Jiménez**

Antes de describir las características, construcción y aplicaciones de una base de datos, es conveniente mencionar algunas definiciones y antecedentes partiendo del *dato*, el elemento básico de esta nueva herramienta de la informática moderna. Se espera que con una secuencia de ideas el lector pueda seguir el tema con más facilidad.

Definición de algunos términos usados en informática

Se considera que un *dato* es el registro de un hecho aislado que es significativo. Al agrupar varios datos sobre un mismo tema y con un criterio regulador, se obtiene *información*. La información ordenada por criterios ayuda a quien la recibe en la toma de decisiones con un fundamento sólido. Si los datos fueron bien tomados y guardan alguna forma de relación entre sí, la información resultante será útil. Por ello, la relación conectiva entre datos se debe establecer con anterioridad a la misma obtención de datos. Los datos son patrones estáticos, que reflejan una realidad, mas no ayudan a interpretarla. La información, en cambio, tiene significado cognoscitivo; está relacionado con el conocimiento de los hechos y es una función cerebral. El dato es almacenable en un archivo o en una máquina electrónica; la información es el producto resultante del dato el cual utiliza el usuario.

El manejo de datos para convertirlos en información es conocido como *procesamiento de datos*. Esa función puede generar, como producto final, una base de datos; ésta puede definirse como una acumulación ordenada de observaciones o anotaciones sobre un tema, hecha de tal manera que "los datos tengan una relación lógica entre sí y una estructura bien definida".

* Presentación hecha en un curso sobre Redacción Técnica celebrado en la Universidad Nacional, Heredia, Costa Rica, del 20 al 22 de noviembre de 1985.

** Ingeniero Agrónomo. Coordinador de INFORAT (Información y Documentación Forestal para América Tropical), Departamento de Recursos Naturales Renovables, DRNR, Centro Agronómico Tropical de Investigación y Enseñanza, CATIE. Turrialba, Costa Rica. 1985.

La base de datos, en términos generales, es el puente entre el dato acumulado y la información; es una forma de convertir materia prima en materia utilizable por el usuario. Otra definición podría ser: la colección de datos relacionados por criterios predefinidos para ser utilizados por un usuario de información. No se trata de un simple archivo de datos sin criterio unificador que hilvane o relacione tales datos. El archivo tiene un solo enfoque utilitarista; en cambio, la base de datos puede tener varios enfoques y en consecuencia servir a varios usuarios con necesidades específicas de información. Esta función de "compartir" datos convertidos en información es característica de la base de datos; así, ésta cruza fronteras dentro del amplio espectro del conocimiento (incluye varias *nociones*).

El archivo es el antecesor inmediato de la base de datos. En su génesis evolutivo, el archivo fue poco a poco sustituido por la base de datos; así, la información archivada fue sistematizada por medio de la computación para facilitar su localización y su recuperación, con un mínimo de esfuerzo y de errores.

Antes de describir la metodología que conduce de la colección de documentos, a su categorización y finalmente, al proceso de computarización, es conveniente hacer una analogía: el manejo de la información --sea por una persona o una institución-- es semejante al sistema nervioso de un organismo viviente; no es el organismo en sí el que genera la información; es su mecanismo interno que controla los impulsos nerviosos. Estos impulsos son comparables con los impulsos eléctricos y ahora con los electrónicos. El organismo, como sistema, reacciona ante un impulso y actúa; finalmente, toma decisiones, sean éstas correctas o erróneas, según su capacidad de reaccionar en forma racional o impulsiva al estímulo sensorial que fue captado y es procesado.

Siguiendo con la analogía se puede decir que en la operación de acceso de datos por enfoque o *nociones* (materias o criterios prefijados), el organismo es el equipo de cómputo (hardware); el impulso electrónico lo da un programa o paquete de programas (software) y el producto obtenido de ese impulso es la información contenida en una bases de datos.

Retrospectiva histórica del procesamiento de información

La humanidad ha reunido información y la ha procesado según sus capacidades de captación, de comprensión y de sus motivaciones. No siempre fueron correctas las interpretaciones o utilizaciones que se dieron a la información por parte de las antiguas civilizaciones; ello condujo a la caída de imperios y a la desaparición de tribus en el Continente Americano. El crecimiento de las poblaciones, cuando no hubo diezma ocasionada por las guerras, el hambre o la peste, ocasionó sobrepoblación y luego, creciente presión económica o mayor competencia entre individuos, hasta llegar a los tiempos que corren en que la supervivencia se hace cada día más difícil. Al aumentar la población, es preciso tomar decisiones cruciales en los ámbitos

gubernamentales y para ello, es necesario manejar o administrar grandes bloques de datos (censos u otras estadísticas) que están interrelacionados por algún criterio "normador". La información generada es supuestamente utilizada para provecho de la humanidad.

Así, se han utilizado tarjetas perforadas o cintas perforadas que transformaron los datos obtenidos en información aprovechable. Estas últimas fueron sustituidas por las cintas magnetóicas o grandes "cassettes" que acumulan y transportan, de un lugar a otro, grandes bloques de información ya ordenada. En los años 70 y ahora en la década de los 80, los equipos y los procedimientos para manejar sistemas de información han alcanzado un desarrollo espectacular.

Las bases de datos como herramientas de difusión de nuevos conocimientos

Inicialmente, fueron los bancos las instituciones que diseñaron sistemas de interconexión informática para centralizar diferentes operaciones y dar servicios múltiples a sus clientes (Ver Figuras 1 y 2). Por esta razón, se llamó inicialmente Banco de Datos a lo que luego se denominó Bases de Datos (Ver Figura 3). *

a. Definición

Ya se definió *base de datos* como una acumulación ordenada de observaciones o anotaciones sobre un tema, hecha de tal manera que "los datos tengan una relación lógica entre sí y una estructura bien definida". Por razones que después se darán, se hace énfasis en el concepto *relación lógica*.

b. Objetivos

De nuevo, algo de historia. El libro y posteriormente, la revista, fueron los primeros vehículos de información. Con el crecimiento descomunal de la cantidad de tales documentos (que son publicaciones primarias) fue necesario producir revistas de índices y compendios (que son publicaciones secundarias) que facilitan la obtención de un determinado documento; pero, poco tiempo después, estos dos elementos fueron insuficientes. Al final de la década de los años 60, para resolver el problema de la explosión de la información, se desarrollaron las bases de datos a través de la automatización. Las bases de datos han sustituido a las publicaciones secundarias. En parte, desempeñan la misma función: ayudan a localizar documentos interesantes.

* En la actualidad, se denomina *Banco de Datos* a las acumulaciones computarizadas de información estadística y *Base de Datos* a los conjuntos de información no estadística.

Como sistema de información que es, la base de datos procura suministrar al usuario datos significativos sobre un tema determinado, en una forma utilizable, en el momento oportuno. Si el usuario sabe utilizar provechosamente esos datos se habrá cumplido con el objetivo básico de la base de datos.

c. Productos

La base de datos ofrece algunos servicios derivados de la consulta hecha por el usuario; ayuda a obtener productos específicos, entre ellos: información impresa, en consultas in situ o bien, en forma diferida (a larga distancia), y documentos originales sobre un tema específico, obtenidos a través de búsquedas selectivas.

d. Usuarios

En el ámbito científico y técnico, los usuarios de las bases de datos son los especialistas en una rama determinada del conocimiento humano que necesitan mantenerse informados acerca de los avances logrados por la investigación. La fuente de información de la base de datos la constituyen los documentos escritos, publicados y distribuidos por las instituciones que se ocupan del quehacer científico. Las necesidades de información de los usuarios son variables que, en América Latina, aún no se han estudiado a fondo; quizás, ya existe algún conocimiento acerca de la identidad del usuario mas no de lo que éste necesita específicamente; todavía no se hacen evaluaciones de lo ya publicado ni de lo que necesitan los investigadores, profesores, estudiantes y otros usuarios de la información científica y técnica.

Propósitos a alcanzar

Se debe definir el propósito básico de un centro de documentación: a quiénes se pretende servir y con cuál grado de precisión. Con el mismo criterio de utilización por los usuarios, la bases de datos debe responder a las necesidades de éstos.

Este criterio determinará el grado de "satisfacción" que se pretende lograr con las consultas que harán los usuarios. ¿Será una capacidad de respuesta amplia o limitada? Si se pretende dar una respuesta muy amplia, la base de datos tendrá una estructura más complicada, entrelazando varias áreas de interés denominadas *campos*, los cuales deben estar separados como unidades independientes pero a la vez entrelazados para que puedan reunirse cuando ello sea necesario y así lograr mayor amplitud de consulta. Sin embargo, es casi imposible diseñar una base de datos con capacidad ilimitada de consulta; de ahí que es necesario fijar con anterioridad cuál será el rango de cobertura que se considera es suficiente para cubrir las necesidades básicas del mayor número de usuarios de una base de datos. Para ello, hay que llegar a una "solución de compromiso" al seleccionar los campos o las áreas de interés (*nociones*) a fin de lograr menos pero más productivas interrelaciones temáticas.

Amplia o reducida en su cobertura temática y en su capacidad de respuesta, la base de datos debe ser periódicamente puesta al día, o sea, darle "mantenimiento" incorporando nuevos documentos a fin de que no quede "obsoleta" demasiado pronto. Antes de que esto suceda se debe generar una nueva base de datos ya ampliada mediante mantenimiento; así, se obtendrá una nueva versión como la nueva edición de un libro.

Necesidad de usar un lenguaje documental internacionalmente aceptado

De cada documento brotan términos de uso frecuente por el especialista que conoce bien su materia temática mas no conoce connotaciones sobre informática. Tales términos constituyen el lenguaje natural y es el utilizado por la comunidad científica que produce una enorme masa de documentos publicados. Pero, para encontrar aquellos documentos en la montaña de publicaciones y así estar en capacidad de evacuar una consulta documental, es necesario convertir los términos del lenguaje natural a otros ya definidos por especialistas en informática: estos segundos términos evitan las vaguedades, redundancias y sesgos en el significado que tiene el lenguaje común. Esta traducción de términos al lenguaje documental facilita la búsqueda de nociones o temas de interés de los documentos y así se llega, con mayor precisión, a entrelazar áreas que interesan a los usuarios. La terminología documental es diferente a la terminología del lenguaje corriente o natural, en los siguientes conceptos:

Lenguaje natural

Comunicación inmediata

simple

conceptos tal cual son

ambiguo

arbitrario

ilógico

libre

Lenguaje documental

Comunicación mediata

compleja

vocabulario "editado"

preciso

controlado

lógico y asociativo

normalizado

Para indizar documentos es necesario traducir los signos del lenguaje natural a los signos o términos del lenguaje documental. La traducción se hace a través de un *thesaurus* o tesoro que es un diccionario a la inversa*. Así, se obtiene un "vocabulario bibliográfico" o "lenguaje documental" que es internacionalmente aceptado. Por ejemplo, el AGROVOC es un vocabulario trilingüe hecho para la terminología agrícola y preparado por la FAO y la Comisión de las Comunidades Europeas (CEE). Es utilizado por el Instituto Interamericano de Cooperación para la Agricultura (IICA). La Figura 4A da una definición del AGROVOC y la Figura 4B presenta tres descriptores tomados

* El tesoro encuentra palabras o términos cuando ya se tiene el significado de los mismos, función opuesta a la del diccionario.

de las listas trilingües.

La aplicación del tesoro asegura la precisión de una base de datos y su utilización internacional en lo que concierne a indización de documentos; además, introduce normalización a las "entradas" y "salidas" de información en las bases de datos. Al utilizar un tesoro confiable se garantiza una comunicación efectiva entre los usuarios de una base de datos; también, permite dar mayor flexibilidad a la base de datos al introducir (o modificar) nuevos elementos temáticos a la lista básica del tesoro. Ya se mencionó que su uso evita las vaguedades e inconsistencias, las cuales reducen la eficiencia de las consultas que se hacen a las bases de datos consultadas.

Amat Noguera, autoridad reconocida en el mundo hispanoparlante en asuntos de técnicas documentales, dice que las *palabras clave* se escogen del lenguaje natural, en tanto que los *descriptores* se seleccionan en el lenguaje documental. Ambos elementos sirven para preparar índices de materias pero su utilización moderna se relaciona con la estructura de las bases de datos. Las Figuras 4C y 4D presentan ejemplos de lenguaje natural y de lenguaje documental.

El uso de descriptores permite que se "compartan áreas comunes" entre diferentes bases de datos al tener la misma fundación en sus estructuras. Así, se universaliza la utilización de las bases de datos y se logra el "cruce de referencias" con lo cual se integran elementos afines en el producto que se obtiene con la base de datos: bibliografías impresas sobre temas específicos. La selección de buenos descriptores ayuda a lograr mayor eficiencia en la operatividad de la base de datos. Estas son algunas sugerencias al respecto: (a) Preferiblemente, utilizar términos salidos de un tesaurus; (b) si no existe el término en el tesaurus, formar uno nuevo, procurando que sea un sustantivo concreto, con significado unívoco, claro y apto dentro de la terminología científica; (c) pensar en el usuario cuando se acuña un nuevo descriptor. El AGROVOC es suficientemente amplio dentro de la temática agropecuaria mas no lo es dentro de la forestal.

El ingreso de datos

Desde un punto de vista teórico (aún no he adquirido experiencia práctica en el desarrollo de bases de datos), trataré de exponer cuál es la filosofía que se ha desarrollado sobre la difícil tarea de ingresar datos a una base de datos.

a. El sistema de manejo de una base de datos

Quando ya se está en posesión de datos confiables (científicamente válidos), se inicia el proceso de ingresar datos por computadora a una base de datos (discúlpeme por la repetición obligada de la palabra *datos*). Para ello, se seguirá un sistema de manejo de datos que escoja, adapte o diseñe un programador (o administrador de bases de datos) que debe trabajar en colaboración con el o los documentalistas que tentan a su cargo los procesos bibliográficos que se deben hacer en una colección de documentos, desde

su comienzo (caracterización de los documentos) hasta haber preparado los formularios de entrada de datos. El documentalista conoce las necesidades de información del usuario. El programador conoce la metodología de la computación: el programa a utilizar (sus capacidades y potencialidad) y las técnicas asociativas de temas a través de descriptores.

El sistema de manejo de datos incluye: el equipo de procesamiento electrónico ("hardware"), el programa que permite tal procesamiento ("software") y el producto final (la base de datos en sí).

Ambos especialistas definen: los campos temáticos que son de mayor interés (con respecto al usuario) y la estructura que sea más conveniente en relación con la asignación de campos y subcampos. También, las unidades de memoria que se debe asignar a cada campo y subcampos. ¿Cuánto espacio de la memoria se asignará a cada campo? ¿Hay suficiente memoria disponible? ¿Cuál es la filosofía del centro de documentación con respecto a la simplicidad o complejidad de la temática de la colección de documentos? O sea, se tiene el propósito de lograr una cobertura amplia o restringida del tema básico? El programador dará la pauta para distribuir la memoria disponible, de acuerdo con la amplitud de los campos, una vez que éstos han sido definidos. Esta tarea no es simple de realizar y todavía no hay en Costa Rica personal adiestrado, a excepción de una pocas personas en la sede de la Dirección General del IICA, en Coronado. En el país hay muchos técnicos en computación mas no en el manejo de información bibliográfica para construir bases de datos.

b. La estructura de la base de datos

Al hacer sus cálculos de unidades de memoria y al definir criterios de estructura y programa a utilizar, el programador debe conocer las características de otras bases de datos que se puedan consultar o a las que se puede enviar información codificada; se debe insistir en que haya compatibilidad entre bases de datos, para que haya enriquecimiento mutuo; la base donante y la receptora de información deben "hablar en el mismo idioma de cómputo".

Es posible entrelazar varias terminales (microcomputadoras que funcionan como terminales) a través de una computadora central que tenga puertos de entrada y de salida, para actuar como mediadora entre distintas micros que operan bases de datos. Así, las consultas (o los envíos) hechos "a control remoto" serán compatibles y por lo tanto, bien aprovechadas. En caso contrario, si los lenguajes de computación son diferentes, se repetirá "el caso de la Torre de Babel", en sentido real y figurado.

c. Campos de interés

En el presente texto se hizo referencia a las *nociones* las cuales son, en realidad, campos de interés. Ambas connotaciones constituyen áreas específicas del conocimiento que se deben integrar a la base de datos mediante descriptores y relacionar en forma jerárquica o asociativa, o bien mediante otras modalidades estructurales (Ver Figuras 5, 6, 7, 8, 9 y 10). Sin embargo, no se deben asignar demasiadas áreas de interés a una base de datos

pues es posible que ésta se "congestione" con demasiada información, lo cual resultará en respuestas erróneas causadas por interconexiones defectuosas o imposibles. Por más fuerte y ágil que sea el programa de recuperación que se utilice, el recargo de datos es -- desde todo punto de vista -- inconveniente.

Las formas de asociación vertical o terminal, en arreglo jerárquico o asociativo, permite cruzar los puentes que separan los campos de interés, en el formato que el programador diseñó. La interrelación entre campos de interés se hace preferentemente a través de las últimas ramificaciones o asociaciones finales, las cuales se representan por medio de descriptores. De aquí la importancia de estos elementos en la construcción de una base de datos.

Ya se mencionó que los campos de interés deben ser unidades separadas pero entrelazadas para efectos de dar respuestas a un amplio rango de consultas. Son como los bastidores de un biombo que están unidos por goznes que permiten que el biombo se cierre o se abra con diferentes ángulos entre cada bastidor. Cada campo admite mantenimiento por separado lo cual es conveniente desde el punto de vista operativo.

Las demasiadas ramificaciones también son inconvenientes. Ocupan demasiada capacidad de memoria lo que reduce la utilización de la base de datos; como los propios campos de interés (de los cuales se derivan) su demasía en número también causa congestión. Otra vez, las "soluciones de compromiso" se imponen. Aquí se repite que es necesario que el programador y el documentalista se pongan de acuerdo previamente en cuanto a la cantidad, estructura y complejidad de los elementos que van a integrarse en una base de datos. Como ya se dijo, esta herramienta de información debe responder a las necesidades específicas de los usuarios, y por ello, ambos especialistas deben trabajar en forma asociada.

Las asociaciones verticales o las terminales (final de una cadena) son más específicas que las asociaciones en red o en matrices, puesto que llegan a mayores profundidades en el significado o connotación del término usado como descriptor.

d. Pruebas de validez de la base de datos

Antes de difundir a otros centros de documentación una base de datos, es necesario someterla a varias pruebas para asegurar su validez e integridad informática. Bajo diferentes criterios de consulta (campos de interés) se deben hacer pruebas de validación e introducir las correcciones del caso. Así, al ponerla en otras manos, se tendrá una mayor certeza de que el producto que se envía tiene una capacidad de respuesta. El proceso de retroalimentación ayudará a detectar fallas de contenido o de estructura, lo cual facilitará la corrección de los errores. Es posible que haya que cambiar los requerimientos de los usuarios, el diseño o la metodología del desarrollo de una base de datos (Ver Figura 11).

e. Limitaciones de las bases de datos

Ya se mencionó que no se debe "congestionar una base de datos con demasiada información ni se debe exceder la capacidad de memoria cuando se está haciendo el ingreso de datos. Si se excede o rebasa esta capacidad, también se producirá un "congestionamiento" de datos y como resultado, la base de datos será defectuosa. Esta es una de las limitaciones en el uso de la base de datos. Otra podría ser la incompatibilidad entre bases por diferencias estructurales.

f. Pruebas de validez de la base de datos

Antes de difundir una base de datos es necesario someterla a varias pruebas para asegurar su validez e integridad informática. Bajo diferentes criterios de consulta (campos de interés) se deben hacer pruebas de validación y procesar las correcciones del caso. Así, al ponerla en otras manos, se tendrá una mayor certeza de que tiene buena capacidad de respuesta. El proceso de retroalimentación ayudará a detectar fallas de contenido o de estructura, lo cual facilitará la corrección de los errores. Es posible que haya que cambiar el diseño o la metodología del desarrollo de una base de datos (Ver Figura 10).

g. Seguridad y respaldo

En ciertos casos, ya no en el ámbito agrobiológico, es posible que la información contenida en una base de datos sea confidencial. En casos de información política o de defensa nacional, la información es restringida o clasificada (secreta). En estas circunstancias, se debe proteger la base de datos cuidadosamente. Entre compañías que manufacturan productos comerciales existe, en cierto grado, un "espionaje industrial" como consecuencia de la lucha por la supremacía en el mercado de productos manufacturados. En el campo científico, quizás, no haya tanta competencia por la posesión de información.

Finalmente, debe existir un buen respaldo ("backup") para la información ingresada en una base de datos a fin de evitar pérdidas de datos. El respaldo que dan las computadoras de gran capacidad de memoria permite la recuperación de información que se ha guardado en discos suaves o cintas magnetofónicas operadas por computadoras de menor capacidad (microcomputadoras).

Las redes computarizadas de bibliotecas

En los países desarrollados ya existen redes computarizadas de bibliotecas que operan como sistemas integrados que sirven a muchas instituciones a través de terminales que reciben y envían información. Una computadora central de gran capacidad sirve como núcleo o base de operaciones y unidad receptora de información procesada fuera del sistema. El ingreso de datos a esta unidad central puede hacerse a través de comunicaciones "en línea" o por medio de satélites. Así, se enriquecen las unidades periféricas de

menor tamaño (en general, microcomputadoras con pantalla, teclado, disquetera e impresora).

El "Ohio College Library Systems, OCLS" sirve a más de 2000 bibliotecas con aproximadamente 6000 terminales. La Universidad de Stanford, en Palo Alto, California, opera un vasto sistema computarizado de terminales que dan acceso a miles de estudiantes quienes consultan a la unidad central por medio de casetas telefónicas distribuidas en el campus y en la ciudad. La inserción de una tarjeta de crédito permite a los muchos usuarios del sistema hacer consultas bibliográficas y obtener documentos impresos.

Algunos países europeos, entre ellos el Reino Unido, Francia, Austria y la República Federal Alemana, tienen sistemas semejantes. En la URSS, está apenas comenzando a utilizar este sistema computarizado de difusión de información. Japón también ha logrado grandes avances.

Los países latinoamericanos o iberoamericanos, como se les quiera denominar, apenas comienzan a generar ideas sobre el particular. Existe una gran dependencia en materia de "hardware" por limitaciones industriales.

Si bien ya se usan las bases de datos en algunos países, como México, su procedencia es extranjera.

Sería muy conveniente que se hagan esfuerzos en el área para producir bases de datos con un contenido autóctono, a fin de que la información que se genere esté adaptada a las necesidades de los usuarios latinoamericanos. Es lógico que esa información también deberá llegar a los países desarrollados para que sea integrada al acervo mundial de información que se capta en los grandes centros de documentación. Así, de tales centros se podrá distribuir a otros centros localizados en países en vías de desarrollo.

Sería deseable que nuestros países integren sus bibliotecas y centros de documentación en módulos semejantes al de Ohio y de Stanford, para citar solamente dos ejemplos de los muchos que ya existen en países desarrollados. Esta integración sería un paso de avanzada en el intercambio científico.

Las conexiones "micro-macro" son factibles de realizar en el área, puesto que el uso de microcomputadoras es cada día más extendido y la tecnología de su utilización está en rápido desarrollo. En la actualidad, hay limitaciones en el uso de "software" con capacidades y funciones específicas.

La situación existente en América Latina en relación con las bases de datos

En la actualidad, nuestros países, en términos generales, están apenas dando los primeros pasos en la modernización de los sistemas computarizados en materia de informática. Si bien existen muchos técnicos en la ciencia de la computación, inclusive en Costa Rica, son poquísimos los profesionales que se adiestran en la computación aplicada a la informática, en particular, en el manejo de referencias bibliográficas y en el desarrollo de bases de datos. Por esta razón, sería sumamente conveniente el que se aproveche

al máximo el personal ya adiestrado en estas labores. Los incipientes centros de documentación han de progresar significativamente para que, en fecha no lejana, puedan nuestras instituciones de enseñanza superior y de investigación contar con información computarizada, debidamente ordenada y disponible a las consultas que se hagan. Ojalá que este objetivo se logre antes de que finalice el presente siglo!

Digno de mención es el esfuerzo que ha hecho el IICA—ya referido en párrafos anteriores— en el campo de la computarización de la información y en el desarrollo de programas o paquetes de operación para producir bases de datos. La tecnología que ha desarrollado este organismo interamericano ya ha sido llevada a Colombia (Centro Internacional de Agricultura Tropical, CIAT) y a Perú (Universidad Agraria La Molina).

TEU

Por otra parte, el CATIE está en proceso de integración de esfuerzos con el IICA y con la Biblioteca Comemorativa Orton, que está localizada en el campus del CATIE, en Turrialba. También, está en marcha la integración con las instituciones de enseñanza superior en Costa Rica, entre ellas, la Universidad Nacional, y con otras existentes en el área centroamericana y del Caribe. Una de la forma de fortalecer los programas de investigación y de enseñanza en los países del área es la generación de bases de datos. El fortalecimiento de tales programas generará más información, la cual, si es bien manejada, difundida y utilizada, contribuirá al logro de un mayor bienestar de los habitantes de las zonas rurales y urbanas en esta porción del Continente Americano.

Agradecimientos

El autor agradece al *Dr. Gerardo Budowski*, Jefe del Departamento de Recursos Naturales Renovables del CATIE, sus comentarios escritos al primer borrador del presente texto y a la *Ing. For. Xinia Robles*, Documentalista de INFORAT, sus comentarios verbales.

Literatura consultada

1. AMAT NOGUERA, N. Técnicas documentales y fuentes de información. Bibliograf S.A. Barcelona, España. 1ª ed. 1979, 485 p.
2. ATRE, S. Data Base: Structural Techniques for Design, Performance and Management. John Wiley and Sons, Inc. New York, NY, USA. 1983, 442 p.
3. CLARK, J.D. Database Selection, Design and Administration. CBS Educational and Professional Publishing. New York, NY, USA. 1980. 233 p.
4. FREILING, M.J. Understanding data bases management. Alfred Publishing Co., Inc. Sherman Oaks, California, USA. 1982. 63 p.

5. JIMENEZ-SAA, H. El uso del microcomputador en bases de datos y otras aplicaciones bibliotecológicas. Presentado en: Primera Reunión de Información Forestal para América Latina y el Caribe, celebrada en Resistencia, Chaco, República Argentina. 9 al 11 de octubre, 1985. (mimeografiado). 16 p.
6. ROSOVE, P.E. Developing computer-based information systems. John Wiley and Sons, Inc. New York, NY, USA. 1967. 584 p.
7. TURNBULL, F. Información bibliográfica retrospectiva automatizada. En: Memorias del Curso sobre Elaboración y Redacción de Artículos Científicos. División de Estudios de Posgrado. Universidad Nacional Autónoma de México. México, D.F. 1984. pp. 11-17.
8. VETTER, M. and MADDISON, R.N. Database Design Methodology. Prentice Hall International, Englewood Cliffs, New Jersey, USA. 1979. 306 p.

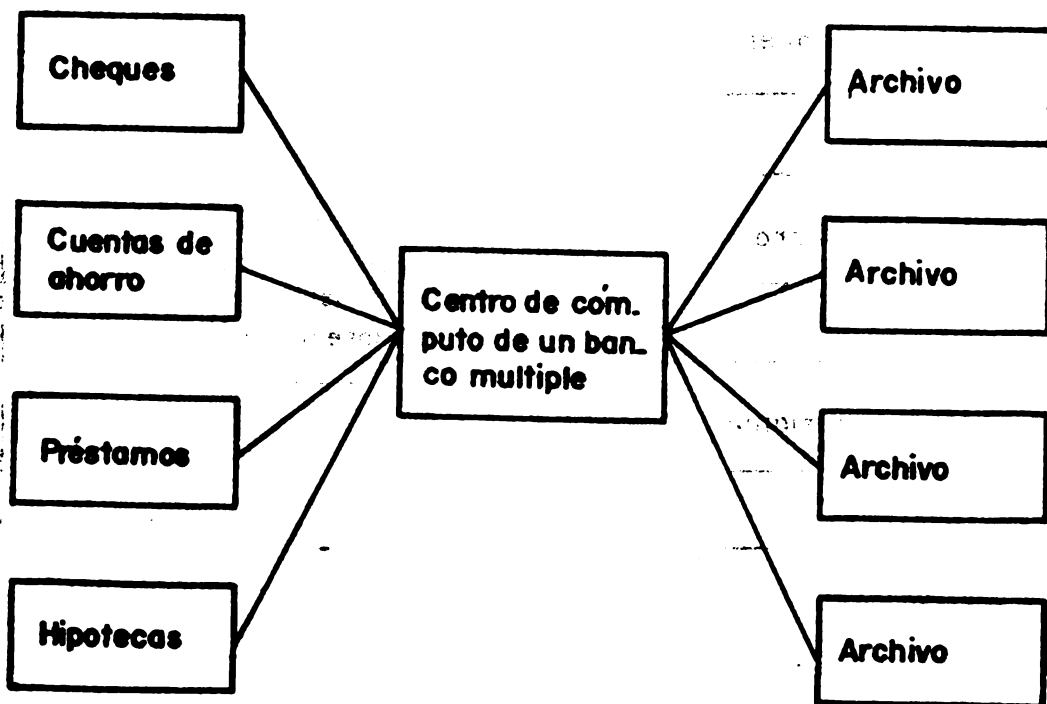
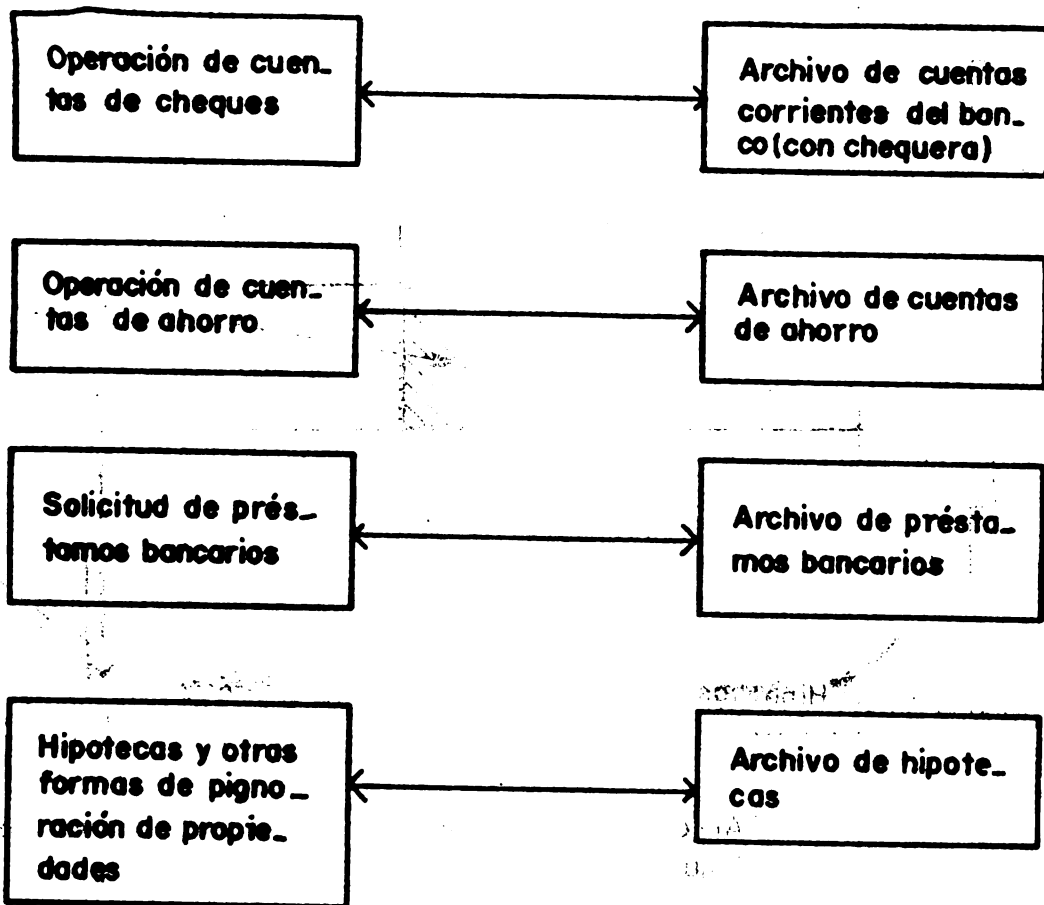


Fig 1. Operaciones bancarias en sus formas clásicas y actual ; esta última evolucionó en un Banco de Datos

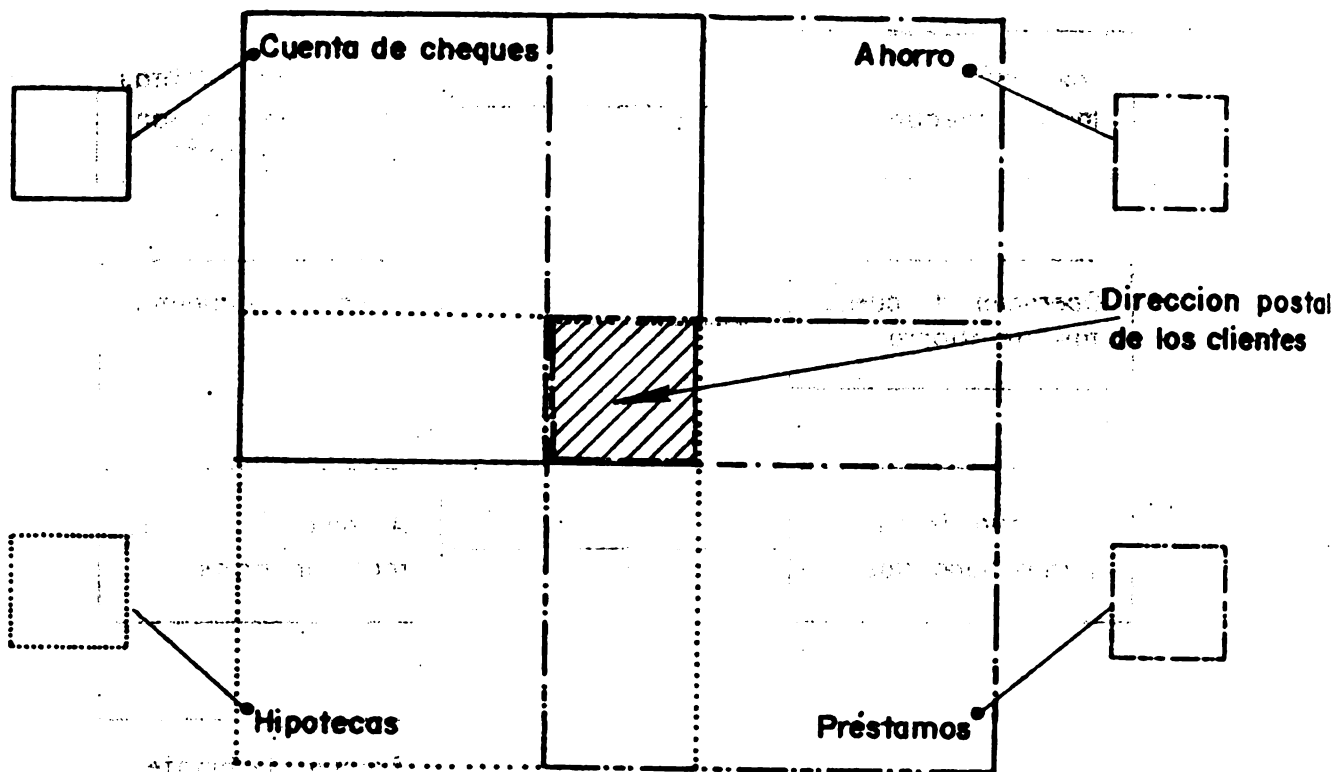


Fig 2. Area común en las cuatro funciones bancarias de la Figura 1.

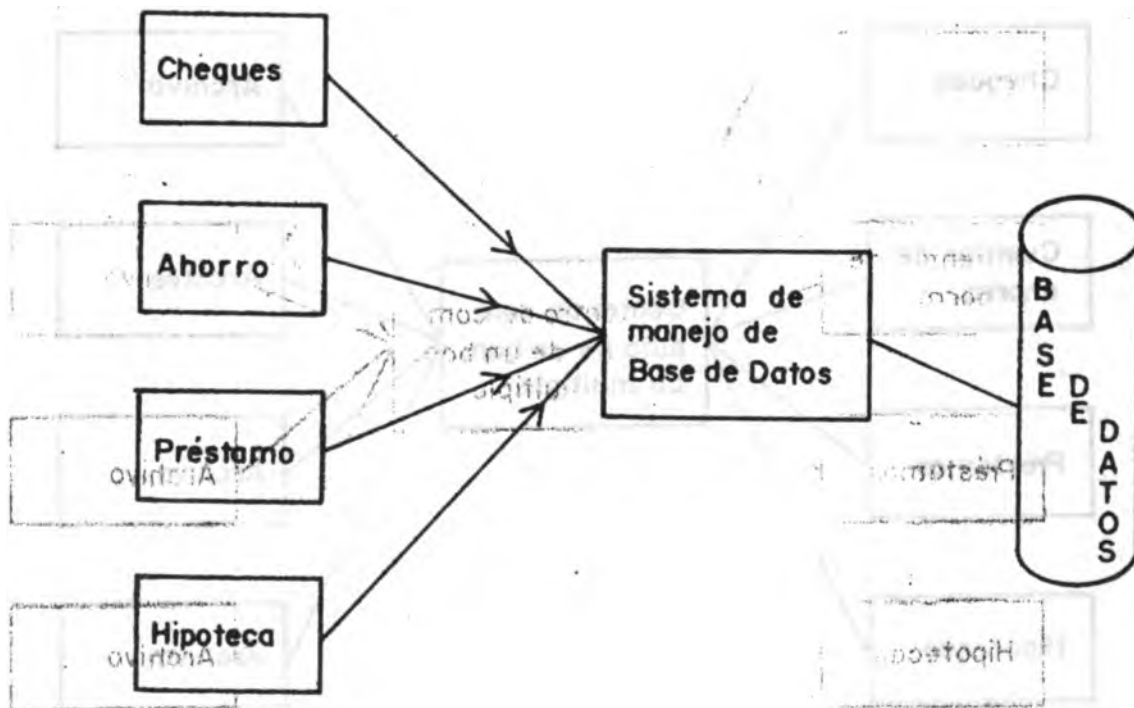


Fig 3. Estructura de una Base de Datos

En el mundo tecnológico de hoy cabe tener acceso instantáneo, mediante computadora, satélite o microficha, a los resultados de las actividades de investigación y desarrollo. Sin embargo, el acceso a esa información depende de un conocimiento común del vocabulario especializado correspondiente a cada materia técnica.

Habida cuenta de la necesidad de material de referencia que hace tiempo se deja sentir, la Comisión de las Comunidades Europeas (CEC) y la FAO decidieron aunar esfuerzos para patrocinar, en abril de 1979, la preparación de AGROVOC, un tesoro multilingüe en alemán, español, francés, inglés e italiano.

AGROVOC, que contiene un vocabulario estructurado de términos agrícolas, tiene por objeto promover una mejor difusión de los resultados de las investigaciones agronómicas y, en particular, permitir una recuperación más eficaz de datos bibliográficos incluidos en el Sistema Internacional de Información sobre Ciencias y Tecnología Agrícolas (AGRIS), dirigido por la FAO.

AGROVOC tiene un alcance mucho más amplio que el de una mera recopilación de términos agronómicos, pues constituye una exposición sintética de toda la materia de la alimentación y la agricultura. Los conceptos básicos y las relaciones entre ellos se expresan con toda claridad en los diversos idiomas y sin prolijas definiciones de diccionario.

Fig 4A. ¿Qué es el AGROVOC ?

Français	ESPAÑOL	English
orycterope abaca lieu jaune alimentation en eau abdomen abeille domestique reine d'abeille	AARDVARK ABACA ABADEJO ABASTECIMIENTO DE AGUA ABDOMEN ABEJA MELIFERA ABEJA REINA	aardvark abaca pollack water supply abdomen honeybees queen bees
Español	ENGLISH	Français
aardvark ácido abscísico abaca oreja de mar terras abandonadas subproductos del matadero mataderos abdomen	AARDVARK ABA ABACA ABALONES ABANDONED LAND ABATTOIR BYPRODUCTS ABATTOIRS ABDOMEN	orycterope acide abscissique abaca onnesau terre abandonnee sous produit d'abattage abattoir abdomen
English	FRANCAIS	Español
abaca offal felling slaughtering kosher slaughter abattoirs abdomen bees honeybees	ABACA ABATS ABATTAGE D'ARBRES ABATTAGE DE BETAİL ABATTAGE KASCHER ABATTOIR ABDOMEN ABEILLE ABEILLE DOMESTIQUE	abaca menudos corta sacrificio matanza kosher mataderos abdomen abejas abeja melifera

AARDVARK

- BT1 tubidentata
- BT2 mamíferos
- BT3 vertebrados
- rt oso hormiguero

← término relacionado usado como descriptor

USE ácido abscísico

← término relacionado usado como descriptor

ABACA

- ul canamo de manila
- BT1 fibras duras
- BT2 fibras vegetales
- BT3 fibras
- BT3 medios de enraizamiento
- BT4 medio de cultivo
- BT3 productos de origen vegetal
- BT4 productos
- rt musa textilis

BT= Broader term (término amplio)

← término relacionado usado como descriptor

Fig 4B. Lista de descriptores tomados del "Índice trilingüe" que es un complemento del tesoro multilingüe de terminología agrícola (AGROVOC) y tres descriptores incluidos en este tesoro.

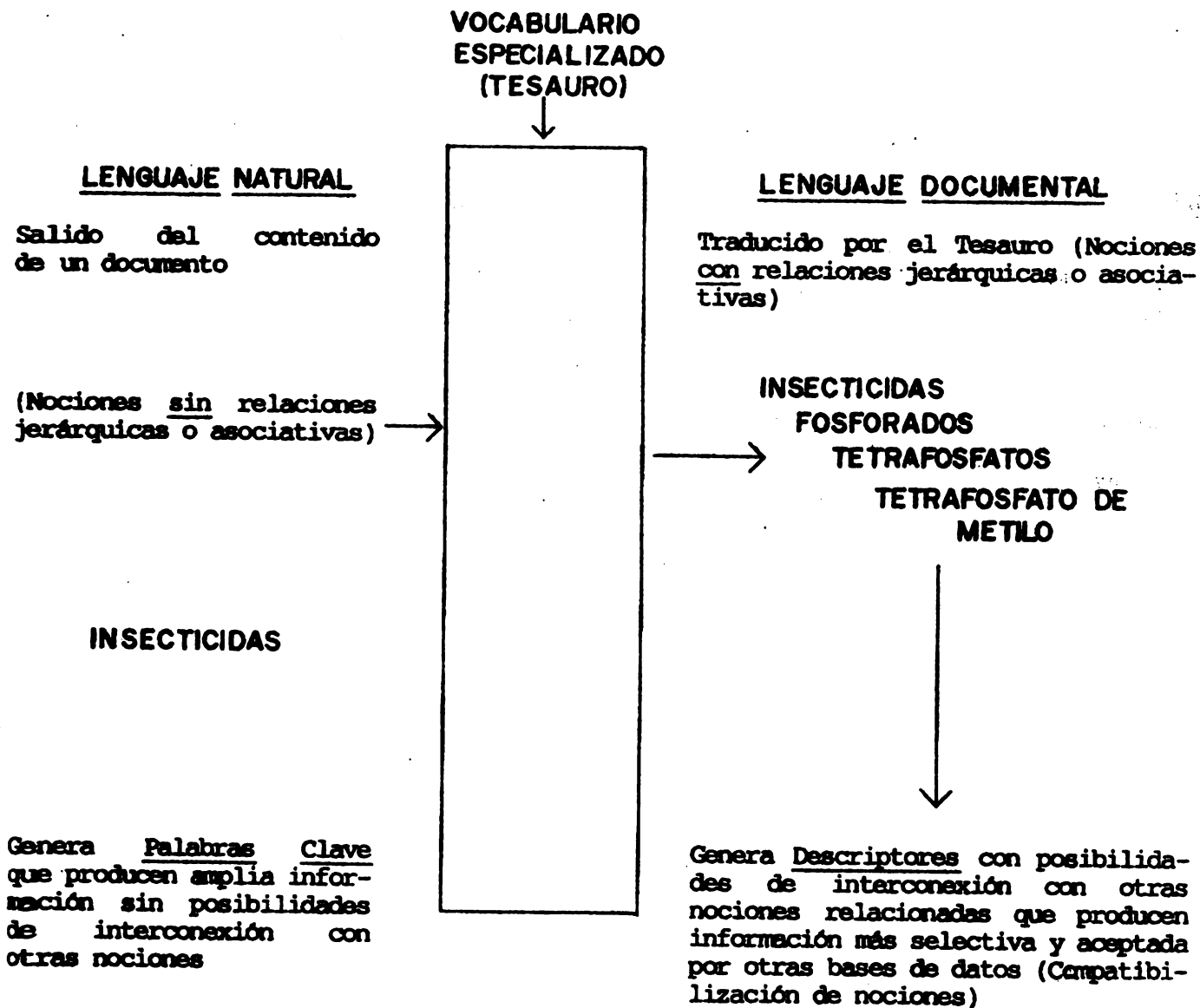


Fig 4C. Diferencia teórica entre palabras clave y descriptor

LENGUAJE NATURAL

Soya

Glicine max

Enfermedades

Hongos

Virus

Bacterias

Cercopora kikuchii

LENGUAJE DOCUMENTAL

GLYCINE MAX

Enfermedades

Fungosas

Virales

Bacteriales

CERCOSPORA KIKUCHII



Descriptores que pueden interrelacionar búsquedas con otras nociones—

PARTES DE LA PLANTA

RAICES

Raíces finas

SUELOS

MAL DRENAJE

Fig. 4D. Ejemplo de interrelación de un descriptor con otras dos (o más) nociones

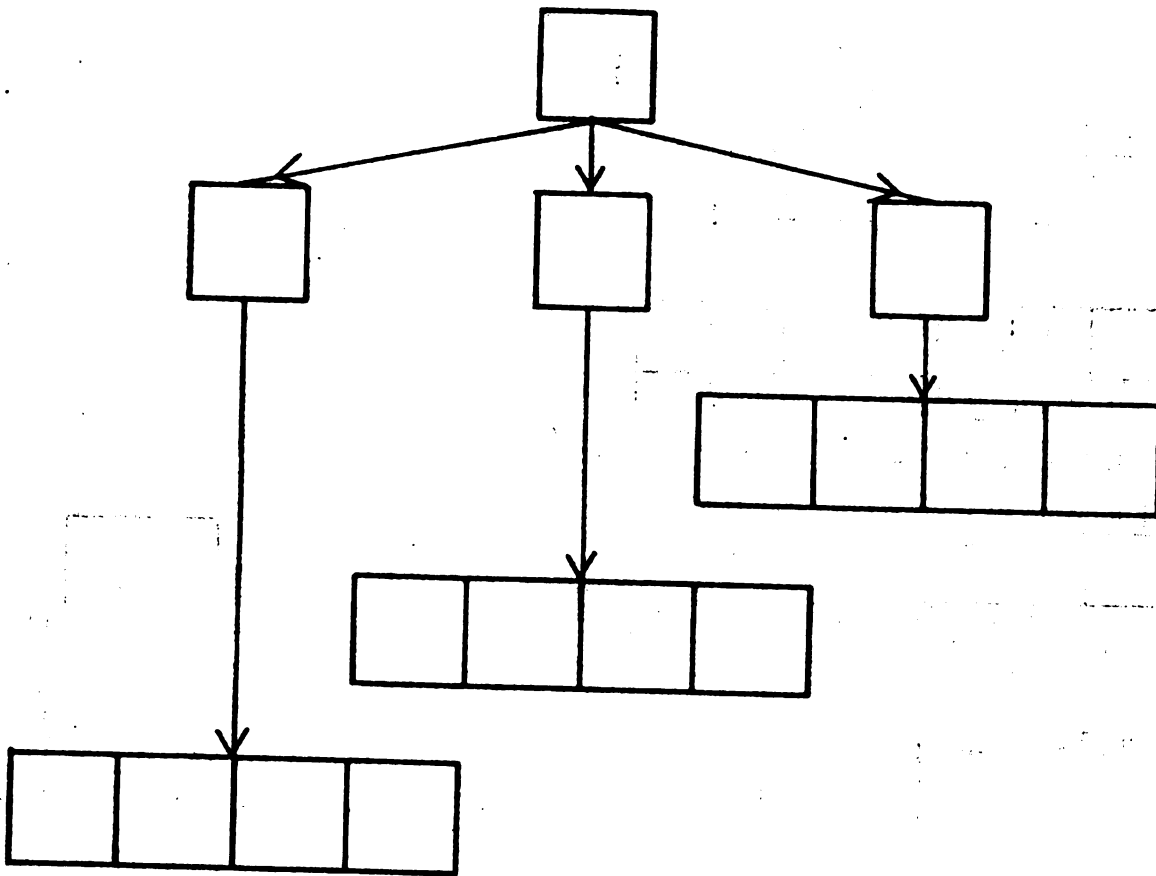


Fig 5. Modelo jerárquico de una Base de Datos

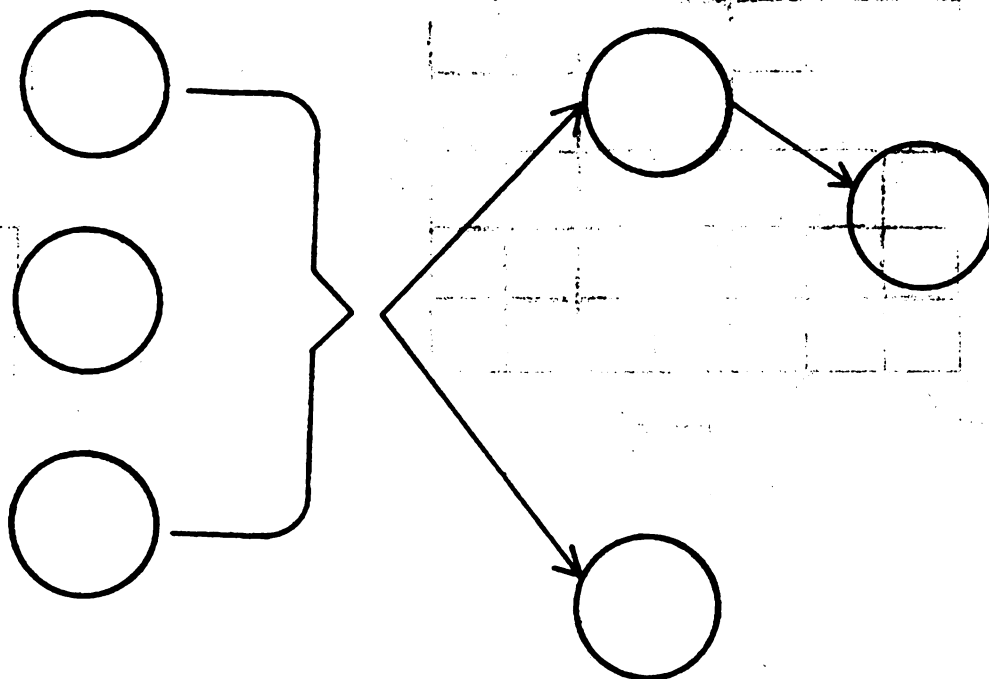


Fig 6. Modelo de red

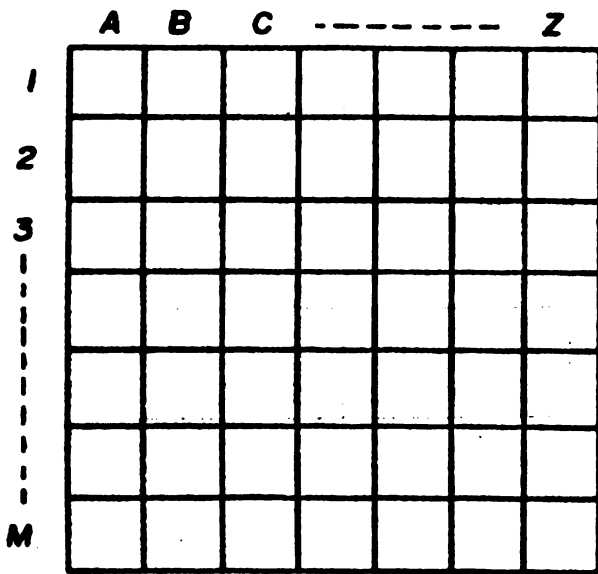


Fig 7: Modelo relacional

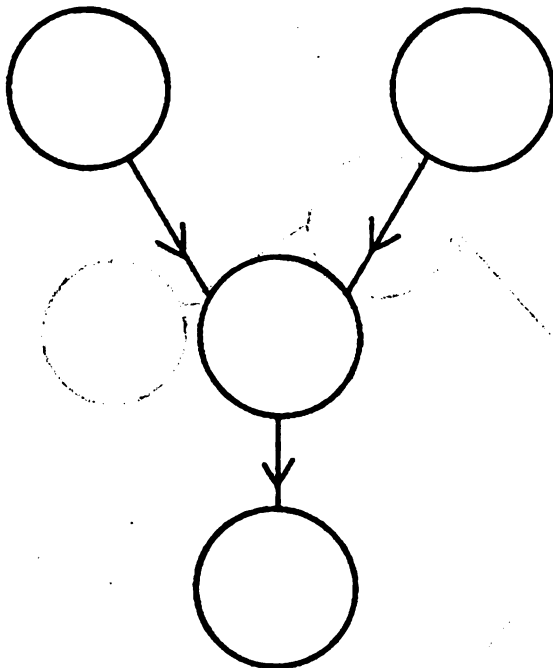


Fig 8. Modelo de Y

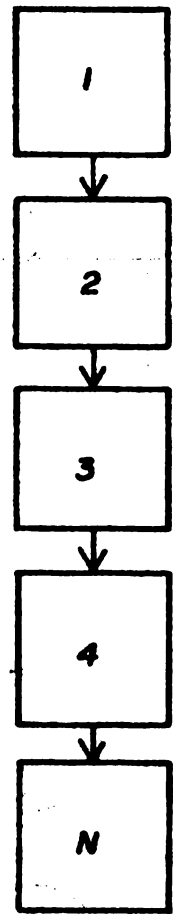
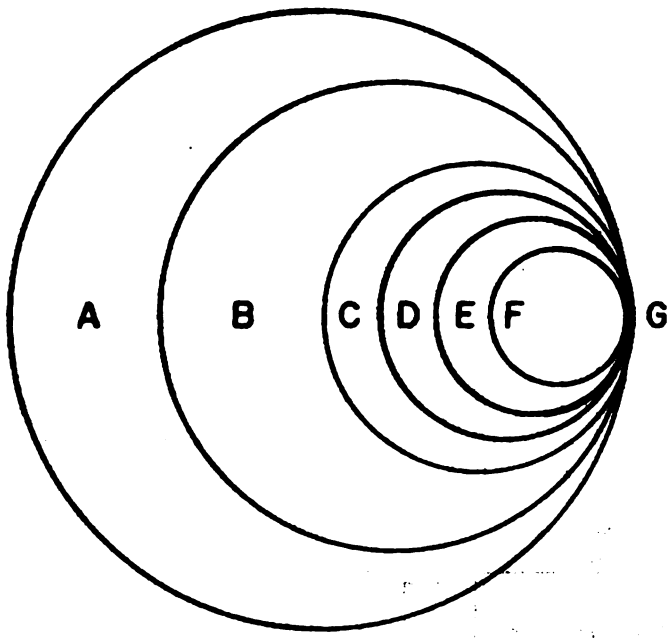


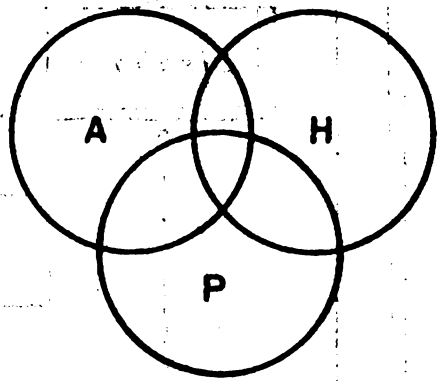
Fig 9 . Modelo de cadena

diagrama
adecuado

*



ORGANIZACION



**a) Estructura jerárquica
(inclusión de categorías)**

**b) Estructura asociativa
(intersección de categorías)**

Fig 10. Estructura de dos estructuras de uso frecuente

Fig 11. Las fases de desarrollo de una Base de Datos

Tiempo Disponible

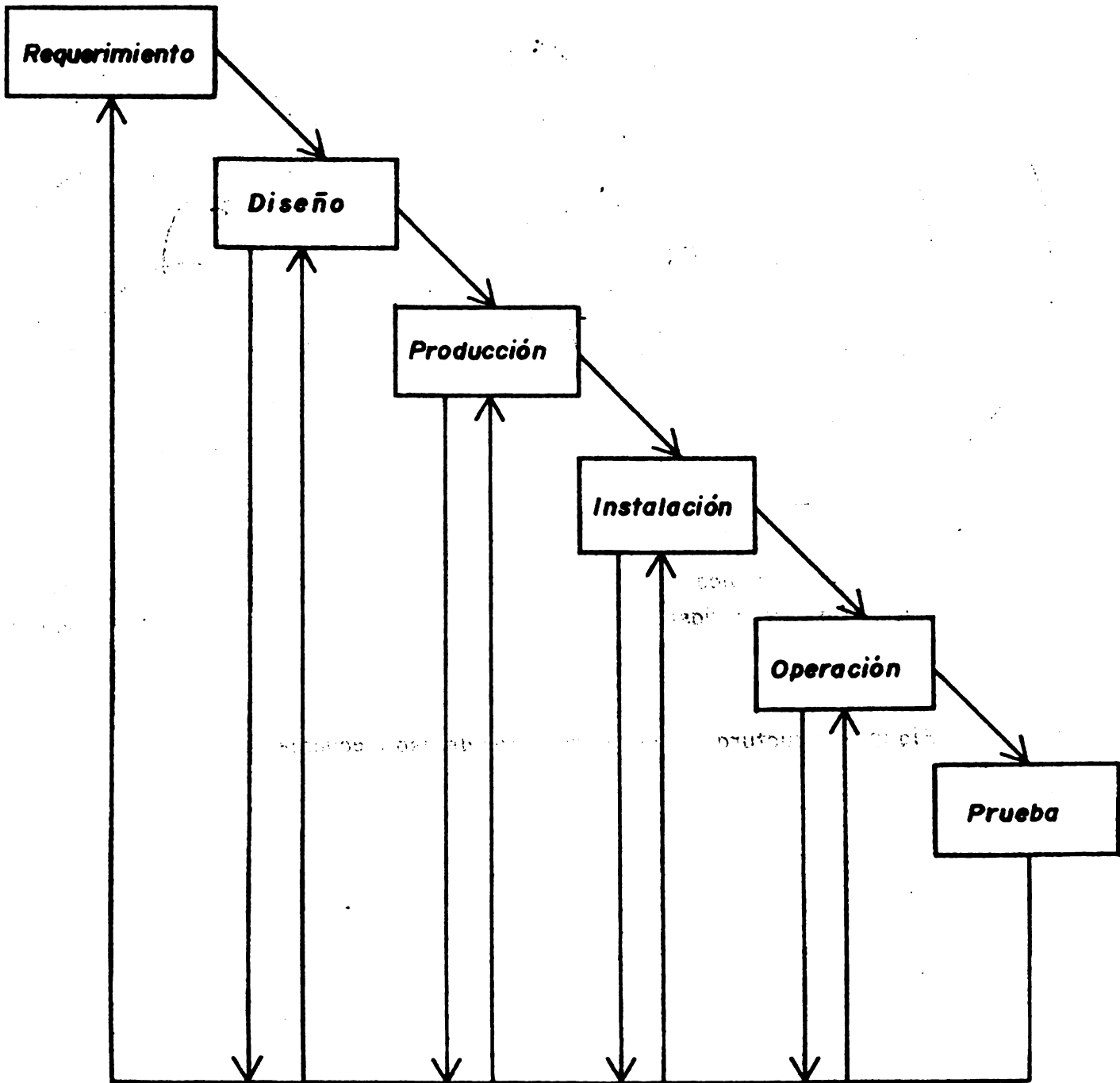


Fig 11. Las fases de desarrollo de una Base de Datos