

ACCESO A LA DOCUMENTACION: BASES DE DATOS

Tomás Baiget

Encargado de información científica. Consorcio de Información y Documentación de Cataluña.

INDICE

1. DATOS, INFORMACION Y DOCUMENTACION. BASES Y BANCOS. TIPOS

2. EVOLUCION DEL ACCESO A LA INFORMACION

2.1 Popularización de la información.

2.2 Edición electrónica

2.3 Telemática

2.4 Soportes ópticos leídos por láser

3. INTEGRACION DE TECNOLOGIAS DE LA INFORMACION

1. DATOS, INFORMACION Y DOCUMENTACION. BASES Y BANCOS

La terminología empleada en Documentación, ciencia y conjunto de técnicas desarrolladas a partir de los años 1940's aún no está completamente establecida y aceptada por todos los que trabajan en ella. Algunas definiciones provocan aún enzarzadas discusiones, agravadas en nuestro país por las discrepancias de criterio de cómo hay que traducir (o no) los términos extranjeros.

A la hora de escribir de estos temas, se hace necesario pues, hacer algunas puntualizaciones sobre algunos conceptos, así como avisar al lector de que no siempre es fácil encontrar términos castellanos que sean fiel traducción del original, generalmente inglés.

Como orientación general, puede decirse que **datos** es la materia prima aportada por sensores, medidores o personas que una vez tratada o elaborada mediante un sistema de información se convierte en **información** útil, lista para ser utilizada para resolver problemas o tomar decisiones.

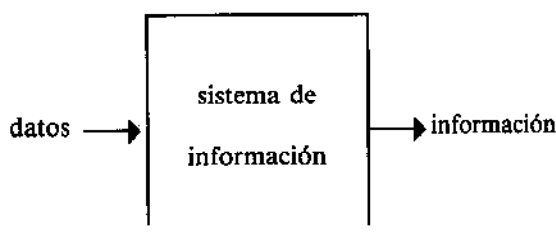


Figura 1. Representación sencilla de un concepto importante.

El concepto de sistema de información, encuadrado dentro de la teoría general de sistemas, ha ganado otra vez mucha importancia, tal como ha sucedido también en otras materias. Por ejemplo, un sistema de información tiene que ser abierto, es decir, capaz de evolucionar y adaptarse a su entorno; tiene que estar realimentado, o sea que tiene que estar evaluado continuamente para que en todo momento ofrezca la información óptima; tiene que ser sensible a un determinado umbral, es decir, para su entrada en el sistema los datos vienen seleccionados por un determinado nivel de calidad o de cantidad, etc. El concepto de entropía, sinónimo de desorden, tiene su equivalente aquí en la falta de información o incertidumbre. Los sistemas abiertos luchan contra la desinformación para hacer disminuir la entropía.

La teoría de sistemas es especialmente importante por su visión global e integradora. Todos los componentes de un sistema funcionan coordinadamente para conseguir los mismos fines. El tratamiento de los problemas de información de una forma conjunta, teniendo en cuenta todos los factores, es necesario no solamente como consecuencia del uso simultáneo de diversas tecnologías, tal como se indica en el apartado 3 de esta comunicación, sino por la complejidad de implantación de las soluciones adoptadas. El técnico de sistemas de información, llamémosle documentalista, debe saber psicología, organización, informática, telecomunicaciones y, por supuesto, documentación.

Dentro de la teoría de sistemas, los sistemas de información son un caso especial puesto que constituyen un subsistema de presencia obligada en todos los otros tipos de sistemas, en cualquier materia.

Cada vez se necesita tratar una mayor cantidad de datos para obtener información, puesto que mayor es la demanda de ella. En las organizaciones, y en la empresa en particular, se va entendiendo que la única forma de funcionar eficazmente, y elevarla productividad para ser competitivos esa base de utilizar mejor la información.

Otro aspecto terminológico a considerar es la difícil delimitación de los conceptos información y documentación. A menudo para obviar la cuestión, y a pesar de la redundancia, se incluyen los dos términos en la misma frase, como sucede, por ejemplo, con el nombre de la institución a la que pertenezco. En general, *información* se refiere al contenido y *documentación* al continente, o sea, al estudio y gestión de los documentos como soporte de la información. En castellano no se ha conseguido encontrar un término equivalente al inglés «information science» que engloba ambos conceptos. Desde algún sector bibliotecario se ha propuesto *biblioteconomía*.

Por último, hay que establecer la diferencia entre bases y bancos de datos, así como entre los varios tipos que hay de ellos. La terminología ha ido evolucionando a lo largo de los últimos 15 años a medida que lo ha hecho la industria de la información ofreciendo online más tipos de información. En algunos casos, se ha tenido que ajustar el significado de los términos a una parcela más específica dentro de las nuevas ofertas de información, y en otros se han tenido que crear términos nuevos.

Banco de datos (tantas veces escrito incorrectamente en mayúsculas —Banco de Datos— por mimetismo del inglés) es el nombre que se dio en un principio a los primeros ordenadores que almacenaban datos a diferencia de los que solamente realizaban cálculos. Esta acepción, que confunde la maquina con la colección o conjunto ordenado de datos, esta completamente en desuso.

En los primeros ficheros informatizados los datos eran buscables de una forma secuencial: para seleccionar las fichas que tenían un término determinado se debía explorar el texto de todas ellas hasta encontrarlas. Posteriormente con el almacenamiento en discos aparecieron los ficheros indizados, en los cuales la búsqueda se realiza de forma mucho más rápida mediante un índice de palabras clave (similar al catálogo de una biblioteca), que remite a las fichas correspondientes. Esta forma de búsqueda, que los bibliotecarios-documentalistas hubieran llamado por índice o catálogo, fue bautizada por los informáticos como «fichero invertido» puesto que para ellos era un procedimiento inverso al secuencial seguido usualmente hasta entonces.

Los ficheros, la mayoría bibliográficos, con entrada a través de índices de palabras clave fueron llamados **bases de datos**, en especial entre los documen-

talistas. Muchos informáticos no están de acuerdo con esta terminología y reservan el nombre a los conjuntos informatizados de datos que están mucho más desagregados y que se pueden relacionar entre sí de cualquier forma con más libertad. Se refieren de forma un tanto despectiva a aquellas bases de datos bibliográficas con la información estructurada de una manera rígida, llamándolas simples ficheros.

Así pues, los ficheros, conjuntos de datos o colecciones de referencias bibliográficas consultados **en línea** (online) siguiendo un diálogo de preguntas y respuestas con el ordenador, se llaman de una forma general **bases de datos**.

En la figura 2 se presenta una clasificación de las bases de datos, en la que se delimita a trazos un grupo constituido por los llamados *bancos de datos*.

Las bdds bibliográficas siempre han constituido el grupo más numeroso y coherente a pesar de las substanciales diferencias que existen entre ellas, en cuanto a la disposición de los campos que integran cada ficha o referencia bibliográfica. Sin embargo, cada vez va más en aumento el número de bancos de datos disponibles. Los bancos de datos son mucho más heterogéneos entre sí y la disposición de los datos depende mucho de los productos o magnitudes que describen o de las ciencias y técnicas a las que dan soporte. Frecuentemente el ordenador ofrece posibilidad de realizar cálculos con los datos almacenados.

La relativa homogeneidad antes mencionada de las bases de datos bibliográficas, se presta más a que sean utilizadas por documentalistas intermediarios que realizan las búsquedas a cuenta de terceros. En cambio los bancos de datos, bien porque requieren un mayor conocimiento de las respectivas técnicas, especialmente los factuales, o bien por el contrario, porque dan una información directamente utilizable y de fácil interpretación (textos completos, catálogos, directorios), tienden a ser más utilizados por los usuarios finales, o sea, los que buscan información para sus propias necesidades.

2. EVOLUCION DEL ACCESO A LA INFORMACION

2.1. Popularización de la información

A medida que la humanidad ha ido avanzando tecnológicamente, la información se ha podido ir «democratizando» llegando cada vez a más personas. En la antigüedad sólo los privilegiados y los poderosos tenían acceso a los papiros y a los pergaminos. La imprenta fue el primer gran paso que permitió la difusión de la información a gran escala. En los años recientes a la informática y las telecomunicaciones han supuesto el segundo gran paso que ha abierto una nueva etapa en la historia.

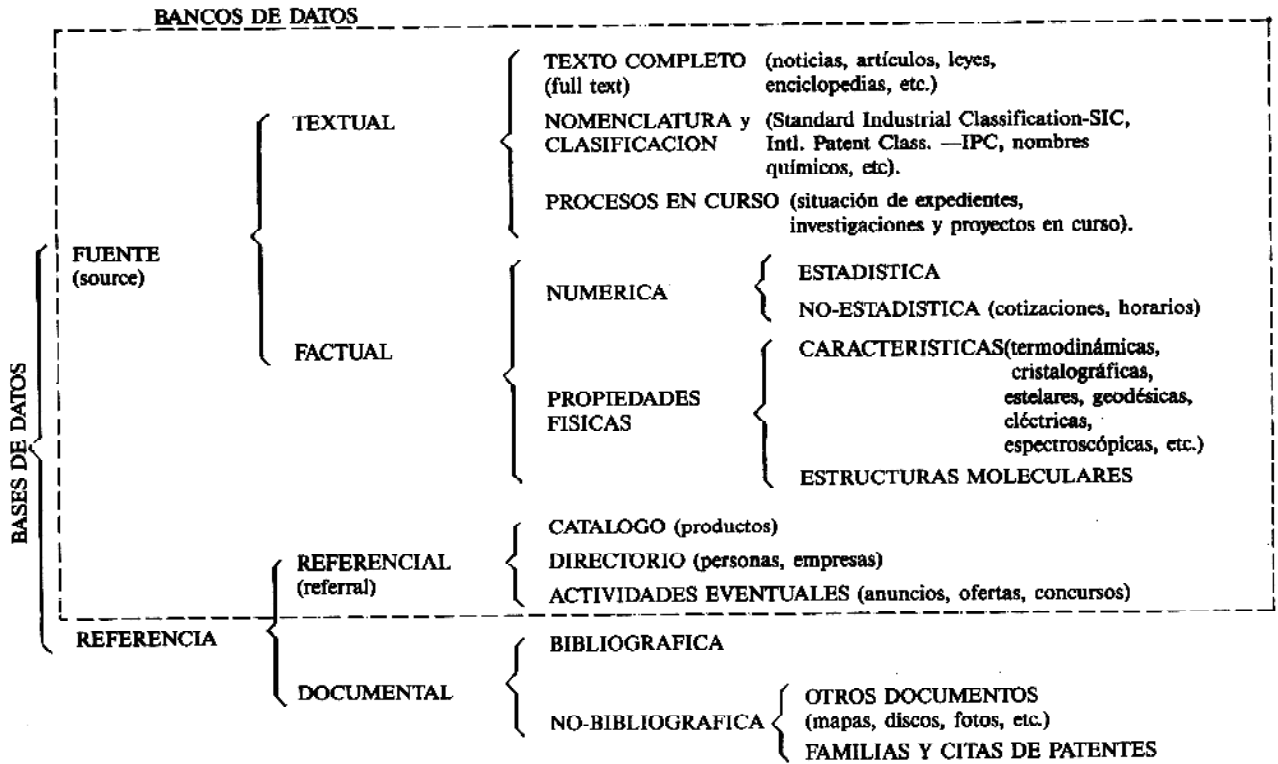


Figura 2. Tipos de bases de datos.

Un aspecto particularmente importante de la popularización de la información es el acceso de los ciudadanos a la información de la Administración pública (ver ref. 1).

En la humanidad se está produciendo un fenómeno de realimentación positiva que lleva a la llamada **explosión de la información**: la información producida se difunde cada vez en mejores condiciones tanto en cantidad como en calidad (rápida, oportuna, fiable,...) de modo que cada vez hay más personas que la usan, los cuales a su vez son capaces de elaborar nuevos estudios o nuevas creaciones, y así sucesivamente. Sin embargo, últimamente se habla ya menos de explosión de la información puesto que con los medios informáticos disponibles se considera que la mayor parte de la información producida, o por lo menos la más importante está ya «bajo control». El problema estriba ahora en como utilizar mejor la ingente cantidad que se tiene fácilmente a mano.

2.2. La edición electrónica

Los primeros ficheros informáticos allá por los 60's eran entrados en los ordenadores a partir de los textos impresos en papel. Más adelante con la posibilidad de entrar los datos a través de una consola o terminal de teclado y pantalla (con facilidades para controlar automáticamente el deletreo, cambiar los formatos, etc.), se pasó a editar primero los ficheros en el ordenador para a partir de aquí sacar la versión im-

presa. La calidad de la impresión ha ido mejorando paulatinamente. Actualmente las máquinas de fotocomposición, alimentadas por las cintas o discos conteniendo los datos, permiten editar los textos con una gran variedad de tipos de letra.

Al soporte papel se le vaticinan aún muchos años de vida. Sin embargo ya existen diversas fuentes de información que sólo se ofrecen por vía electrónica, para ser leídas por una pantalla. La sustitución del papel por la pantalla parece utópica a algunas personas, que hablan de la «sensualidad» insustituible de tener un libro o una revista entre las manos. Se está aún en el período más balbuciente de la futura era de las pantallas, pero el excepticismo de algunos frente a las mismas ha empezado a desaparecer ya con algunos modelos que tienen una imagen muy fija, gran definición y son de lectura tan confortable como el papel. Se avanza en la reducción del tamaño de las mismas y existen pantallas muy planas (2 cm). Todo es cuestión de saber aprovechar los nuevos servicios de información que se pueden obtener a través de una pantalla, y de habituarse a ellas. No se trata de perder el placer de los libros sino de saber hallar uno nuevo en la edición electrónica. Vale más que no parapetemos nuestras mentes contra ellas porque llegan de forma imparable.

Antes se ha hecho referencia a la popularización de la información. Otra cualidad que las nuevas tecnologías de la información, en especial gracias a las telecomunicaciones, han traído consigo, es su univer-

salidad, su ubicuidad. Es muy sencillo tener acceso desde casa a los catálogos de la Library of Congress, o a la bolsa de valores de Londres. También es fácil para todos los investigadores y técnicos del mundo poder compartir el mismo **tanque de conocimientos**, con el mismo nivel del **estado de la técnica** y poder integrarse en el mismo **frente de investigación** con un mínimo decalaje de tiempo. Esta facilidad da lugar a una gran acción dinamizadora sobre la investigación.

2.3. Telemática

Podría hablarse aquí de las diferentes técnicas y servicios modernos que tratan u ofrecen información mediante la integración de la informática y las telecomunicaciones, en un contexto social, encuadradas bajo el nombre genérico de telemática. En la siguiente tabla se enumeran los principales:

Teledocumentación	Sistema de acceso interactivo (en línea) a bases de datos (distantes), mediante un terminal (asíncrono), normalmente a través de las redes (públicas) de transmisión de datos.
Telefacsimil	Transmisión de fotocopias. En la actualidad una página A4 tarda menos de 1 minuto (equipos de los grupos 3 y 4).
Videotex	Sistema de teledocumentación, que en Europa distribuyen las propias compañías telefónicas, que como pantalla terminal usa el televisor doméstico. La información está organizada en forma de menú y el usuario indica las opciones escogidas mediante una especie de mando a distancia. Pueden conectarse varios ordenadores proveedores de información a un mismo sistema de videotex mediante gateways . El sistema de videotex español de CTNE se llama Ibertex.
Teletext	Sistema en el cual la información es emitida página a página desde una emisora de televisión simultáneamente a la programación normal. Los usuarios la reciben en casa a través de la antena y mediante una especie de mando a distancia pueden seleccionar las páginas deseadas para que aparezan por la pantalla del televisor. La emisora envía continuamente todo el ciclo de páginas (unos centenares en total) en intervalos de muy pocos segundos. El aspecto de las pantallas es parecido a las del videotex pero el teletext tiene mucha menos capacidad en número de páginas y no es interactivo.

Teletex

Sistema de télex muy perfeccionado que usa terminales inteligentes. Los mensajes, que pueden componerse fácilmente en local mediante un programa de tratamiento de textos, pueden guardarse en memoria enviarse a varios usuarios simultáneamente, etc., a gran velocidad. Usa terminales especiales determinados por las compañías telefónicas de cada país.

Correo electrónico

Nombre que genéricamente puede referirse a cualquier sistema de envío de mensajes por medios electrónicos, pero que cada vez se va reservando más a los sistemas que usan terminales asíncronos estándar. Muchos ordenadores que ofrecen servicios de teledocumentación dan también servicio de correo electrónico, con características parecidas al teletex.

Tablón de anuncios electrónico y teleconferencia

Sistema de tormo electrónico con «buzones» compartidos por varios usuarios que pueden sólo leer, o leer y escribir. Pueden organizarse «reuniones» sobre temas determinados durante unos días o permanentemente (club de debates).

Ofimática

Conjunto de nuevas tecnologías de información aplicadas al trabajo en la oficina. En una primera etapa la automatización de la oficina sólo ha sido para facilitar tareas aisladas, pero poco a poco con nuevos servicios antes inexistentes, se va yendo a un cambio profundo en la forma de trabajo, en la comunicación, en la toma de decisiones y en la forma de solucionar los problemas, basándose en un uso eficiente de los recursos de información. Dentro de la ofimática se estudia: tratamiento de textos, creación de bases de datos, redes de telecomunicaciones locales (para compartir equipos, instalar correo electrónico, monitorizar procesos, etc.) y otros servicios telemáticos.

2.4. Soportes ópticos leídos por láser

Sin embargo, se hace obligatorio referirse aquí a las tecnologías de almacenamiento óptico que en 1986 han representado un enorme impacto en el campo de la documentación, especialmente el CD ROM (discos compactos de memoria leible y no borrrable).

CD ROM (Compact Disk Read Only Memory). Grabación digital (bits), 4.75" = 12 cm. de diámetro. Imagen, sonido y texto. El usuario no puede grabarlo. Pista espiral de más de 5 Km. Capacidad 600 Mbytes/cara (grabado con caracteres alfanuméricos tiene una capacidad de 180.000 páginas A4, y por imágenes facsímil 7.500 páginas).

DISCO DIGITAL (disco óptico digital, disco óptico numérico). Grabación digital (bits), 12" = 30 cm. de diámetro. Imagen, sonido y texto, digitalizados mediante scanner. El usuario puede grabarlo pero no borrarlo (WORM = write once read many). capacidad 1.000 Mbytes/cara (grabado con caracteres alfanuméricos tiene una capacidad de 500.000 páginas A4, y por imágenes facsímil 30.000 páginas). Marcas: Megadoc (Philips), Gigadisc (Alcatel Thomson).

VIDEODISCO (disco óptico analógico, vídeo LP). Grabación analógica (surcos de diferente grosor), 12" = 30 cm. de diámetro. Imagen, sonido (y texto). El usuario no puede grabarlo. Capacidad 54.000 frames o imágenes de televisión (audio + vídeo), todas direccionables, por lo que también se llama videodisco interactivo. Poca resolución para textos de letra pequeña.

Todos estos Sistemas llevaban varios años desarrollados pero sólo ha sido el CD ROM, gracias a su estandarización consecuencia de un convenio entre Philips y Sony, el que se ha comercializado de forma masiva.

Los discos compactos, aún en plena fase de introducción suponen una nueva forma de distribuir (o de tener acceso a) la información, que compite con la teledocumentación y las versiones impresas o revistas de resúmenes. A primera vista pareció que la más afectada sería la primera pero ahora, se ve que los usuarios más lógicos son los actuales suscriptores de las versiones impresas, es decir, instituciones que a cambio de una cantidad fija de suscripción anual tienen acceso ilimitado a la base de

datos. Si la teledocumentación ha ido paulatinamente restando clientes de las versiones impresas, se espera que el disco compacto lo haga mucho más. Los productores de bases de datos habían tenido que ir encareciendo las revistas de resúmenes y aumentar los cánones (royalties) de consulta en línea para mantener el nivel de ingresos. El disco compacto es mucho más barato de producir y de transportar que la revista se compra también por suscripción anual (cada dos o tres meses se recibe un nuevo disco con la base de datos actualizada, teniendo que devolver el anterior), y permite la consulta de la base de datos como si fuere en línea pero con la ventaja de que no hay que pagar por el tiempo de búsqueda. Como consecuencia, se les augura muy pocos años de vida a las bases de datos en soporte papel.

DISTRIBUIDOR DE BDDS

PRODUCTOR DE BASES DE DATOS

CINTA MAGNETICA

HOST

ACCESO EN LINEA
(teledocumentación)

CARGA Y CONSULTA EN ORDENADOR PROPIO

DISCO COMPACTO

ABSTRACTS

REVISTA DE RESUMENES

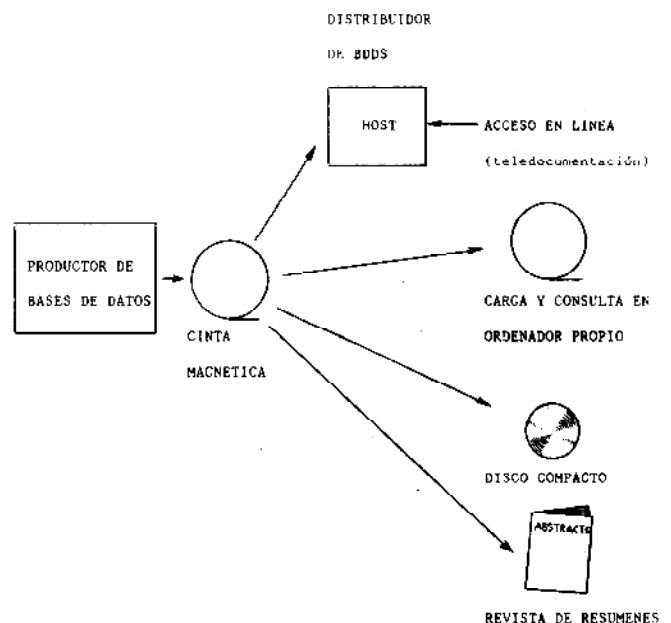


Figura 3. Formas de distribuir las bases de datos.

Los grandes usuarios de determinadas bases de datos podrán comprarlas en disco para consultarlas en local, pero tendrán que seguir como clientes de los sistemas en línea para consultar los millares de otras bdds disponibles de uso menos frecuente.

3. INTEGRACION DE TECNOLOGIAS DE LA INFORMACION

Tal como ya se ha dicho en el apartado 2.3. referido a telemática y, concretamente, en el subapartado de informática las tecnologías de la información se van integrando gradualmente dando lugar a nuevos servicios. Algunas veces una nueva aplicación se consigue con una tecnología realmente nueva pero otros muchas veces, se trata de combinaciones de equipos de, por ejemplo, tamaño más reducido, nuevo diseño, nuevo formato o nuevas prestaciones, gracias a los cuales esa aplicación empieza a ser cómoda o rentable.

Las tecnologías de la información se basan en tres grandes puntales: tecnología de oficina, comunicaciones e informática. Todos ellos evolucionan muy rápidamente, lo cual hace la tarea del diseño de sistemas de información muy arriesgada puesto que lo que hoy es una buena solución, fruto de un estudio serio de un problema concreto, puede dejar de serlo en pocos días por la aparición en el mercado de un nuevo equipo mejor y más barato.

En la figura 4 se indican las diferentes tecnologías de la información a partir de los tres componentes señalados anteriormente. Por su importancia cabría subdividir la informática en dos: proceso y almacenamiento.

INFORMATICA
(proceso+almacenamiento)

ORDENADORES
Y SUS PERIFERICOS

teletratamiento

terminales portátiles

conmutación de paquetes

redes inteligentes (VAN, LAN)

informática distribuida

gateways

(interconexión de hosts)

teletext

videotex

teledocumentación

tele-entrada de datos

activación por voz

transferencia
electrónica de fondos

teletex

correo electrónico

teleconferencia

bases de datos

sistemas expertos

discos ópticos/
almacenamiento multimedia

editoras de sobremesa

tratamiento de textos

copiadoras programables
e inteligentes

COM/CIM

tarjeta perforadora

télex

teledictáfono

telefacsimil

TELEFONOS,
CENTRALITAS

COMUNICACIONES

MAQ. DE ESCRIBIR,
COPIADORAS

TECNOLOGIA DE OFICINA

Algunas tecnologías ya se han comentado en el capítulo anterior. Resaltar unas u otras depende más bien del entorno en que nos movamos. Lo más importante a destacar, es la tendencia a su uso conjunto por parte de un mismo usuario de forma cotidiana. Es decir, todas ellas han pasado o pasarán pronto, directa o indirectamente a influenciar el funcionamiento de las empresas y nuestra forma de trabajo, nuestros hábitos y nuestro ocio.

En relación a las bases de datos, título de esta comunicación, en la figura 4 se detallan explícitamente sobre la línea informática-tecnología de oficina. Se incluyen aquí, tanto las bases de gestión (nómina, contabilidad, etc), como las de información que se indican en la figura 2.

Gracias a la disponibilidad de potentes microordenadores muchas organizaciones se plantean la posibilidad de informatizar sus archivos o colecciones de documentos para convertirlos en bases de datos. Esta no es una tarea fácil, que casi siempre se subestima y subvalora, dando lugar a fracasos demasiado frecuentes. La general ignorancia que existe sobre documentad hace que

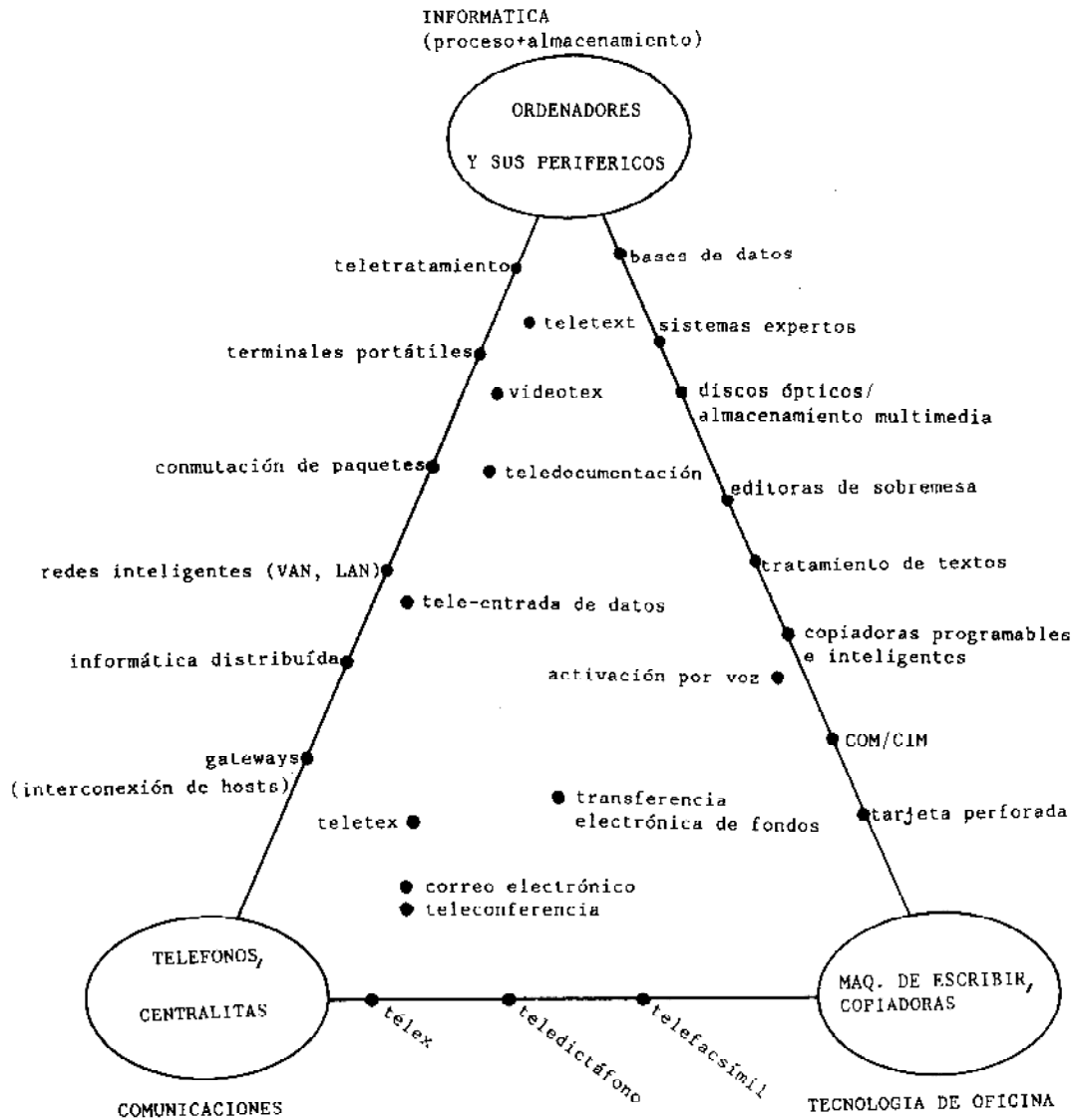


Figura 4. Integración de tecnologías de la información.

BIBLIOGRAFIA

1. Bravo, Joan. El acceso a los documentos de la Administración pública. La documentación y la informática como instrumentos de perfeccionamiento democrático. I Congreso Iberoamericano de Informática y Documentación. Medellín (Colombia) Nov. 1985. Temas y Documentos. Núm. 96. Secr. Gral. Técnica. Ministerio de Cultura. Madrid. Febrero-marzo 86.
2. Mason, Donald. Information management. Peter Peregrinus Ltd. Stevenage, UK. 1978. p. 121.
3. Tapscott, Don. Office automation. A user-driven method. Plenum Press. New York & London. 1982. p. 244.
4. Baiget, T. y Meléndez, V. Información online 1985-86. En: Segundas Jornadas Españolas de Documentación Automatizada. Torremolinos. 20-22 Nov. 86. pgs. 15-26.
5. Rispa Márquez, Raúl. La revolución de la información. Aula Abierta. Salvat núm. 99. Madrid. 1982. p. 64.

muchos piensan que con una máquina y un programa de creación y recuperación se tenga ya una base de datos, sin contar con el trabajo de clasificación e indización de los documentos.

El acceso a bases de datos externas (teledocumentación, videotex) es una tarea mucho más sencilla y «agradecida» puesto que se puede tener un servicio eficiente con un esfuerzo económico y personal relativamente moderado.