

FACULTAD DE FILOSOFÍA Y LETRAS
COLEGIO DE BIBLIOTECOLOGÍA

*“DESARROLLO DE BÚSQUEDAS EFECTIVAS DE
INFORMACIÓN EN BASES DE DATOS DEL CENTRO
ELECTRÓNICO DE INFORMACIÓN DEL INSTITUTO
TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE
MONTERREY, CAMPUS CIUDAD DE MÉXICO”*

TESIS PROFESIONAL QUE PARA OBTENER EL TÍTULO
DE LICENCIADA EN BIBLIOTECOLOGÍA

PRESENTA:

MARTHA ORALIA GUERRERO CAMACHO

289886

ASESORA: BEATRIZ CASA TIRAO

MÉXICO, D.F. 2001



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

Agradezco infinitamente a DIOS por haberme dado la oportunidad de Vivir lo que he vivido y de Conocer y Amar a las personas que amo pero sobre todo de amarlo a Él.

Gracias a la Maestra Beatriz Casa, por ser un ejemplo vivo a seguir, me inspira un profundo respeto, le agradezco haberme ayudado y guiado, los conocimientos que con gusto me transmitió y su calidez son cosas que nunca olvidaré...

Con todo mi cariño esta tesis se la dedico a mis padres, a mi hija y a mi esposo.

A mis padres: Celia y Benjamín, por ser el mejor ejemplo, agradezco los valores y la confianza que siempre me dieron, gracias por ser mis Padres, a ustedes les debo lo que soy.

A TI MADAI POR SER LA FUERZA QUE ME MUEVE DÍA CON DÍA Y ME IMPULSA A SUPERARME, ERES LO MEJOR QUE DIOS ME ENCOMENDÓ.

A Alejandro, en especial por su apoyo incondicional, compañía y amor que hicieron posible este momento, gracias.

Con cariño a mis hermanas Celia, Lucía y Susana, por ser mis eternas compañeras; a Reyna, Chuy y Tití que hicieron posible todo, a Pacesita y a mis Abuelos Benjamín y Jorge, Q.E.P.D.

A mis profesores y a la UNAM por que me brindaron todo para poder llegar a este momento.

Al Sistema ITESM y al personal de Informática y de la Biblioteca del CCM por que me siento orgullosa de ser parte del equipo.

A toda mi familia entera, a mis amigos, compañeros y equipo de trabajo y todas aquellas personas que contribuyeron en alguna forma para la realización de este trabajo y que permanentemente me estimularon para el logro de este objetivo personal.

CONTENIDO

Páginas

Introducción

Capítulo 1. INTRODUCCIÓN A LAS BASES DE DATOS

1.1 Definición y tipología.....	1
1.1.1 Bases de datos referenciales.....	4
1.1.1.1 Directorios.....	6
1.1.2 Bases de datos factuales o numéricas.....	7
1.1.3 Bases de datos textuales.....	8
1.1.4 Bases de datos mixtas.....	8
1.1.5 Otros criterios de clasificación.....	9
1.1.5.1 Por el área temática.....	9
1.1.5.2 Por el ámbito territorial.....	10
1.1.5.3 Por su acceso o procedencia.....	10
1.1.5.4 Por su origen.....	11
1.2 Modos de acceso a las Bases de datos.....	11
1.2.1 Conexión en línea.....	11
1.2.2 Acceso local.....	13

Capítulo 2. RECUPERACIÓN DE INFORMACIÓN EN BASES DE DATOS

2.1 La información y su importancia.....	19
2.2 La recuperación de la información.....	22
2.3 Recuperación de la información en bases de datos.....	24

2.3.1 Tipos de recuperación de la información en las bases de datos.....	28
2.3.2 Indicadores de la eficacia de la recuperación.....	31
2.4 Modelo de métodos y técnicas de recuperación de información.....	33
2.5 Lenguajes informáticos, métodos y técnicas.....	34
2.5.1 Operadores boléanos y lógicos.....	35
2.5.2 Operadores de expansión.....	38
2.5.3 Operadores de intervalo.....	39
2.5.4 Operadores de proximidad o distancia.....	39
2.5.5 Operadores de cualificación.....	40
2.5.6 Operadores de encadenamiento.....	41
2.5.7 Otros recursos y procedimientos de exhaustividad.....	42

Capítulo 3. ESTRATEGIAS DE BÚSQUEDA

3.1 Cómo determinar el tema específico a buscar.....	44
3.1.1 Definir un tema específico para la búsqueda.....	45
3.1.2 Seleccionar los términos de búsqueda.....	45
3.2 Preguntas más importantes al comenzar una investigación.....	48
3.3 Método para instrumentar una estrategia de búsqueda efectiva.....	51
3.3.1 Proceso para desarrollar una búsqueda.....	51
3.4 Selección de las bases de datos a utilizar.....	53
3.4.1 Variaciones de las bases de datos.....	53
3.4.2 Por qué elegir entre una y otra base de datos.....	54
3.4.3 Sugerencias a tener en cuenta para seleccionar una base de	

datos.....	55
------------	----

Capítulo 4. EL INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

4.1 El Sistema ITESM.....	57
4.1.1 Misión del ITESM.....	58
4.2 El Campus Ciudad de México del ITESM.....	60
4.3 La Biblioteca del ITESM CCM.....	61
4.3.1 Datos generales de la biblioteca del ITESM CCM.....	62
4.3.2 Misión de la Biblioteca.....	63
4.3.3 Objetivos de la biblioteca.....	63
4.3.4 Políticas generales de la biblioteca.....	64
4.4 El Centro Electrónico de Información (CEI).....	65
4.4.1 Misión del CEI.....	65
4.4.2 Objetivos del CEI.....	66
4.4.3 Recursos del CEI.....	66
4.4.4 Servicios del CEI.....	67
4.4.5 Usuarios del CEI.....	67
4.4.6 Ventajas del CEI.....	67
4.4.7 Descripción de las principales bases de datos en línea del	69
4.4.8 Descripción de las principales bases de datos en disco compacto del CEI.....	76
4.4.9 Descripción de las principales bases de datos instaladas en	

disco duro en el CEI.....81

4.5 Selección de la base de datos para la aplicación del cuestionario.....83

Capítulo 5. DIAGNÓSTICO DEL RESULTADO DE LAS BÚSQUEDAS EN BASES DE DATOS DEL CEI

5.1 Experiencia a priori.....100

5.2 Cuestionario.....101

5.3 Evaluación de la aplicación del cuestionario.....104

Capítulo 6. PROPUESTA PARA FACILITAR EL EMPLEO ÓPTIMO DE LAS TÉCNICAS DE BÚSQUEDA DE INFORMACIÓN EN BASES DE DATOS DEL CEI

6.1 Factores a considerar para la elaboración del curso.....108

6.2 Planeación del curso de capacitación de “Introducción a las bases de datos y la búsqueda y recuperación de la información en bases de datos.....110

6.2.1 Antecedentes.....110

6.2.2 Estructuración del curso.....111

6.2.3 Objetivos generales.....112

6.2.4 Programa.....112

6.2.5 Técnica.....114

6.2.6 Material de apoyo.....114

6.3 Contenido del curso de capacitación.....115

6.4 Selección de dos grupos: experimental y de control.....116

6.5 Aplicación del curso en el grupo experimental.....120

6.5.1 Resultados.....	120
6.6 Comparación de los dos grupos.....	121
6.6.1 Análisis comparativo de los dos grupos.....	127
CONCLUSIONES Y RECOMENDACIONES.....	130
OBRAS CONSULTADAS.....	134
LISTA DE SIGLAS.....	137
ANEXOS.....	138

INTRODUCCIÓN

El tema que he elegido como investigación para mi trabajo de tesis es, sin duda, un tema muy complejo pues la falta de bibliografía en México y la carencia de estudios e investigaciones, son limitantes, sin embargo, para nosotros: bibliotecarios y profesionales que estamos relacionadas con el ámbito informativo y documental considero que nos debe preocupar y debemos abordarlo e investigar, pues se está posponiendo algo que es sumamente importante difundir. Ejemplo, la información en las nuevas tecnologías: las bases de datos y la recuperación de información a través de éstas.

“En un mundo globalizado, la información juega un papel fundamental y ésta es apoyada de manera importante por los recursos de la tecnología y la información...como las nuevas tecnologías no son ajenos al proceso de globalización sino que forman parte consubstancial del mismo”¹.

La importancia de las bases de datos radica en la información que contienen y los métodos de recuperación de las mismas. El acceso a la información facilita a los usuarios e investigadores su crecimiento dado que les permite conocer las noticias actuales, los sucesos importantes, el mercado, los avances tecnológicos, etc.

¹ Casa Tirao, Beatriz. “La información en un mundo globalizado”. En *Revista Relaciones Internacionales*. No. 71, jul-sep., 1996. p. 99.

Las estrategias de búsqueda permiten a los usuarios hacer un uso eficiente de los recursos de la institución, ya que pueden orientar mejor sus esfuerzos y obtener mejores resultados de manera más rápida.

En este trabajo, el objetivo es que a través de la investigación se pueda mostrar lo importante que es actualmente saber manejar las nuevas tecnologías de información, en particular las bases de datos y el desarrollo de estrategias de búsquedas de información. Teniendo en cuenta, que si al usuario se le capacita en este ámbito, es posible que estas estrategias ayuden de manera considerable a mejorar el tiempo y sus esfuerzos al hacer una búsqueda. Para sustentar esta propuesta se desarrolló el proyecto de planeación y ejecución del curso, en el cual se consideraron aspectos que resultan fundamentales para ser conocidos y realizados en una investigación con recursos de información electrónica. En el curso, sin duda, quedaron plasmados tanto los conocimientos que se obtuvieron de la investigación documental, como la experiencia que se adquirió a través del trabajo en el Centro Electrónico de Información.

Desarrollar el programa para el curso, elaborar las presentaciones de apoyo, aplicarlo y recoger finalmente los resultados, ha representado una experiencia hondamente gratificante y enriquecedora en cuanto mi formación profesional y espero también que sea una aportación que coadyuve a resolver las cuestiones relacionadas con el uso efectivo de las tecnologías de información en las Bibliotecas o Centros de Información de cualquier Institución.

Pienso que es un tema no explotado en México y debemos darle sentido profesional, pues así el bibliotecario actual se convertirá no sólo en prestador de información pasivo, sino en, un "facilitador" de la información.

Primeramente se realizó una revisión documental para conocer todo lo relacionado con las bases de datos, la búsqueda de información electrónica y las estrategias que se emplean comúnmente con fuentes de información electrónicas para tener los conocimientos y profundizar en el tema a manera de saber todo lo relacionado en este ámbito.

La presente tesis se divide en siete capítulos, en el primero se desarrolla de manera teórica lo concerniente a las bases de datos desde su definición, los diversos tipos de bases de datos que existen hasta las diferencias en el acceso. Con un panorama general se puede conocer cuál es la situación actual con la información que se encuentra en formato electrónico y saber las características y las diferentes alternativas que brindan, al igual que sus especificaciones particulares.

En el capítulo dos, titulado: "Recuperación de información en bases de datos", se abordan diferentes temas que a su vez sirven de reflexión al lector como la información y su importancia, los diferentes tipos de recuperación de información que hay, los tan importantes operadores booleanos así como otros operadores que no se conoce tan ampliamente como estos últimos y el lenguaje que se requiere para consultar una base de datos y una parte de indicadores en la eficacia de la recuperación de información que nos sirven para estar más guiados en la

búsqueda y en cualquier momento poder determinar de manera objetiva si nuestra búsqueda fue o no exitosa.

Teniendo en cuenta lo anterior, en el siguiente capítulo se trata lo relacionado a las estrategias de búsqueda, claramente se sugieren diversos cuestionamientos que se puede hacer cualquier usuario en el momento que necesita realizar una búsqueda de información y saber de manera general cómo seleccionar desde el tema y los términos de búsqueda hasta un método completo para instrumentar una estrategia de búsqueda efectiva en bases de datos.

Para conocer los antecedentes en donde se lleva a cabo la investigación en el capítulo cuarto de la tesis, se describe al Sistema ITESM desde sus inicios así como el Campus Ciudad de México hasta la Biblioteca y obviamente el Centro Electrónico de Información, lo interesante de este capítulo es que al realizar la descripción del CEI se hace referencia a su colección y es en este momento cuando se enlistan las diferentes bases de datos que existen en el Centro, desde las que su formato es en línea, en disco compacto y las que se instalan en disco duro; después de esto se integra a este capítulo la selección de la base de datos para realizar la investigación.

Después de esto, y con el ánimo de hacer una investigación completa se hace un diagnóstico previo para saber cómo se encuentran de manera general los usuarios, si están o no acostumbrados a buscar información en fuentes electrónicas y qué tan eficaces son con este tipo de formato; para esto se aplicó un cuestionario y se pudo ver de manera sencilla pero contundente sus necesidades y las razones por las que no recurren a este tipo de información.

Ya para finalizar la investigación, se realizó una propuesta para facilitar el empleo óptimo de las técnicas de búsqueda de información en las bases de datos del CEI; los factores que se debían considerar, la planeación del curso de capacitación, el contenido, etc. y lo más importante de esto fue la selección de los dos grupos con los que se llevaría a cabo la investigación, el grupo experimental y de control, se compararon los dos grupos y con esto se pudo observar de manera objetiva y precisa cuáles fueron los resultados y las diferencias que tienen cuando a un usuario se le capacita en el manejo de las bases de datos y la búsqueda y recuperación de información pues se hizo un análisis comparativo de los dos grupos y se dejaron ver las bondades y las ventajas que tiene esta capacitación.

Por último se exponen las conclusiones finales y se hacen algunas recomendaciones para poder servir de referencia al ITESM, específicamente a la comunidad del Campus Ciudad de México.

CAPÍTULO 1

INTRODUCCIÓN A LAS BASES DE DATOS

La información que se utiliza con fines académicos se encuentra en diversos tipos cuyas características obedecen al contenido de la información presentada y al uso para el que son destinados. Constituyen tipos diferentes de fuentes de información, independientemente del soporte en que se encuentran presentadas: papel, discos ópticos, medios magnéticos, etc. Las bases de datos son actualmente una herramienta útil e importante para el ámbito informativo, por lo que al igual que otras fuentes de información, merece ser estudiadas y analizadas.

A continuación describen algunas características generales relativas a las bases de datos.

1.1 DEFINICIÓN Y TIPOLOGÍA

En general, se comienza por dar una definición, pero en esta oportunidad comenzaré por decir qué NO es una base de datos.

Una base de datos no es lo mismo que el programa informático que se utiliza para explotarla.

Aunque muchas veces se oye y se lee la expresión "base de datos" referida a los programas, parece más apropiado reservar a estos últimos el término genérico de "sistemas de gestión de bases de datos" SGBD o DBMS *-Data Base Management System-* en su versión inglesa y en consecuencia, utilizar "base de datos" para

al conjunto de la información almacenada. No obstante, está claro que se trata de conceptos muy relacionados, en la medida en que sería poco operativo tener una serie de datos sin un programa que los maneje, o programa sin datos que manejar, lo cual significa que ambos conceptos interactúan.

Especificando las definiciones, son muchas las que es posible encontrar. Acabo de apuntar que conviene reservar la expresión de bases de datos para la información almacenada y al respecto se puede decir la siguiente:

“Base de datos: colección de datos almacenados electrónicamente en un formato predefinido y acorde con un conjunto de reglas preestablecido; a menudo se conoce por esquema”¹

“Una base de datos es una colección de archivos interrelacionados creados con un DBMS: Data Base Management Systems. El contenido de una base se obtiene combinando datos de todas las diferentes fuentes en una organización, de tal manera que los datos estén disponibles a todos los usuarios, y los datos redundantes puedan eliminarse o al menos minimizarse.”²

Otra definición de bases de datos es la siguiente:

“Una base de datos puede definirse como una colección de datos interrelacionados, almacenados en conjunto sin redundancias perjudiciales o innecesarias; su finalidad

¹ Held, Gilbert. Diccionario de tecnología de las comunicaciones. Madrid: John Wiley, 1997. p. 60.

² Tsai, Alice. Sistemas de base de datos administración y uso. México: Prentice-Hall Hispanoamericana, 1973. P. 46.

es la de servir a una o más aplicaciones de la mejor manera posible; los datos se almacenan de modo que resulten independientes de los programas que los usan; se emplean métodos bien determinados para incluir datos nuevos y modificar o extraer los datos almacenados. Dícese que un sistema comprende una colección de bases de datos cuando éstas son totalmente independientes desde el punto de vista estructural.”³

Así, después de un análisis de diferentes definiciones se elige la siguiente:

“Una base de datos es cualquier conjunto de información almacenada en cualquier tipo de soporte de modo que sea "amigable" y “legible” por un sistema informático.”⁴

Habitualmente, el modo de explotar una base de datos consiste en seleccionar del conjunto total un subconjunto que reúna una serie de condiciones determinadas por el usuario.

La mayoría de las definiciones son similares a la anterior pues tratan acerca de la información a la que se accedía mediante terminales remotas.

Al margen de esto, de la definición escogida me interesa resaltar los términos de “amigable” y “legible”. De la calidad de la información almacenada y de las posibilidades y recursos del programa informático que se emplee, dependerá en gran medida que tal proceso de selección se realice rápida y eficazmente. Esto es, que se recupere toda la información requerida y sólo ella. En cualquier caso, y para plantear una idea más cabal del asunto, merece la pena abordar el tema de los diferentes tipos

³ Martin, James. Organización de las bases de datos. México: Prentice-Hall Hispanoamericana, 1987. p. 57.

⁴ De Akin, Rose. Bases de datos: uso y administración en centros de cómputo. México: Trillas, 1988. p. 14-15.

de bases de datos.

Hay que señalar que existe actualmente una gran variedad de tipologías y clasificaciones de las bases de datos. El criterio más usual es el que establece la distinción atendiendo al contenido de la información almacenada. Desde este punto de vista, en una primera instancia suele distinguirse entre bases de datos referenciales y bases de datos de fuentes primarias; las últimas contienen todo el texto o los datos originales mientras que las primeras remiten a la fuente original, y de la misma sólo incorporan una referencia.⁵ Además, las primeras almacenan documentos secundarios (referencias) y las segundas documentos primarios (originales, fuentes).

Por lo mismo propongo la siguiente clasificación:

- ◇ Bases de datos referenciales: bases de datos bibliográficas, directorios.
- ◇ Bases de datos fuentes: bancos de datos factuales o numéricos
- ◇ Bases de datos textuales.
- ◇ Bases de datos mixtas.

1.1.1 BASES DE DATOS REFERENCIALES

Se trata de uno de los tipos de datos más frecuentes. Las fuentes originales suelen ser elementos impresos (artículos de revistas habitualmente), aunque también los documentos primarios pueden estar en otro tipo de soportes (material audiovisual, programas informáticos, etc.). Lo importante es que del documento original, en la base

⁵ Huang, Xiangji. "Design and Implementation of a Chinese full-text retrieval system based on a probabilistic model". En: Computer, Communication, Control and Power Engineering. Conference: IEEE, 1993. Vol.2 p: 1090 -1093

de datos, almacena sólo información de referencia. Por ejemplo, para el caso más usual de artículos de revistas científicas, la información almacenada suele responder al siguiente esquema:

- Autor o autores del artículo.
- Título del mismo.
- Nombre y datos (fecha, número, páginas) de la revista donde ha sido publicado.
- Palabras-clave o descriptores que pretenden describir el contenido del artículo.
- Resumen del mismo.

En este caso la información se estructura de tal manera que cada uno de estos apartados constituye un campo, y el conjunto de todos ellos configura el registro relativo al documento en cuestión.

A su vez, el conjunto de todos los registros conforma la base de datos. El esquema reseñado es un ejemplo sencillo. La mayoría de las bases de datos incorporan más campos, como más se verá adelante.

Este tipo de bases de datos se denomina bibliográficas porque siguen el esquema habitual del tratamiento de este tipo de información aunque, insisto, no todos los documentos originales estén publicados o su soporte sea en papel. Como antes se mencionó, existen bases de datos sobre películas, fotografías, grabaciones sonoras y material audiovisual en general. Lo mismo ocurre con los programas informáticos. Aunque lógicamente los campos varían en función del tipo de documento primario, el esquema es siempre similar. El procedimiento que se sigue para integrar esta base de

datos es el siguiente:

El documento original es analizado por un documentalista, y como resultado de este análisis se incorpora a la base de datos la referencia bibliográfica del mismo.

Este conjunto de información estructurado en campos y que conforma el registro de cada documento tiene una doble finalidad. Por un lado, hay campos que se emplean prioritariamente para seleccionar la información requerida: las palabras-clave o la fecha de publicación, por ejemplo. Otros, por el contrario, se incluyen en el registro para informar al usuario sobre el contenido y características del documento original. Es el caso del resumen "abstract", en la terminología inglesa. Dado que mediante la consulta a la base de datos no puede accederse al documento original, su función estriba en que el usuario pueda, mediante su lectura, evaluar con más conocimiento de causa, si los artículos previamente seleccionados responden o no a la demanda de información efectuada. Si esta evaluación de la pertinencia es positiva, el usuario, con el resto de los datos, puede localizar los documentos originales.⁶

1.1.1.1 DIRECTORIOS

Sobre la base del mismo esquema, en este caso la información almacenada hace referencia habitualmente a datos relativos a personas o instituciones.

Los casos más frecuentes son los *Who's who* y los directorios de empresas, organizaciones e instituciones públicas normalmente agrupados por áreas de actividad.

⁶ Sherrod, John. "Information systems and networks". En: Eleventh annual symposium. march 27-29, 1974. Westport: Greenwood. 1975. p. 95.

1.1.2 BASES DE DATOS FACTUALES O NUMÉRICAS

Aunque se emplean indistintamente ambos calificativos, lo importante es señalar que en este caso la base de datos contiene información numérica de diversa índole y organizada según diferentes criterios.

Entre otros ejemplos se puede citar los siguientes resultados de encuestas de opinión pública, indicadores económicos, cotizaciones de bolsa, estadísticas oficiales, censos, etc. En este tipo de bases de datos la explotación suele ser de dos formas:

En la primera se interroga al sistema en demanda de una información muy concreta y, por tanto, la recuperación de la misma debe realizarse considerando un elevado número de condiciones. Por ejemplo: número de habitantes de sexo masculino de un pueblo determinado, mayores de 18 años según el último censo realizado, otro ejemplo podría ser la producción anual de trigo de determinada comarca. El segundo tipo de consulta es la que busca seleccionar un conjunto grande de datos con vistas a un tratamiento posterior, normalmente estadístico, por ejemplo, evolución de los resultados electorales de una demarcación determinada.

Para concretar, en las obras relativas a las aplicaciones telemáticas no es difícil encontrar la siguiente distinción entre ambos términos, el de "banco de datos" que es el lugar donde se encuentra la computadora central que almacena y distribuye toda la información (host), cada uno de los diferentes ficheros en los que se incorpora constituye una "base de datos" independientemente del tipo de información que contenga. Por tanto, de ahora en adelante, al hablar de "bancos de datos" se referirá a una base de datos que almacena información de tipo numérico, exclusiva o mayoritariamente.

1.1.3 BASES DE DATOS TEXTUALES

También se denominan de texto completo o integral. En este caso, el documento original completo se incorpora a la base de datos, de tal manera que cada registro está formado no sólo por la referencia bibliográfica del documento original, sino por el texto íntegro de dicho documento.

Los ejemplos más frecuentes de este tipo de bases de datos se encontrarán en aquellos campos donde al usuario no le basta con la información referencial suministrada por las bases de datos bibliográficas y, por el contrario, requiere todo el texto del documento original. Hay más áreas de aplicación como luego se abordará, pero ahora se citará este caso, el jurídico-legal.

Por ejemplo, el caso de un juez o abogado interesado en conocer las sentencias dictadas en los últimos años sobre un asunto cualquiera. Está claro que, independientemente de la importancia que en cada ordenamiento jurídico tenga la jurisprudencia, un sistema que se limite a informarle de los números de dichas sentencias y añada un breve resumen de las mismas, es bastante incompleto. Lo normal es que requiera la lectura del texto íntegro pues la información que busca probablemente se encuentre en tal o cual considerando o en un determinado apartado del texto.

De ahí que la edición de colecciones jurídicas y legislativas en CD-ROM haya tenido un espectacular avance. No en vano uno de los mayores problemas que plantea este tipo de bases de datos es que requieren mucha capacidad de almacenamiento

1.1.4 BASES DE DATOS MIXTAS

Por último, existen casos de bases que mezclan información de ambos tipos (textual y numérica). La distinción entre los tipos de bases de datos establecidos se funda en la

pregunta de ¿qué tipo de información es la que mayoritariamente incorporan?; ya sea textual o numérica.

Es por esto, por lo que al hablar de las bases de datos textuales decía que la información en ellas contenida no es o no era sólo de tipo referencial.

1.1.5 OTROS CRITERIOS DE CLASIFICACIÓN

Para finalizar este apartado dedicado a los tipos de bases de datos, a través de la experiencia obtenida con el tiempo en el uso y manejo de las bases de datos; se establecen las siguientes tipologías atendiendo a otro tipo de criterios:

1.1.5.1 POR EL ÁREA TEMÁTICA. Esto tiene que ver con lo relativo a lo que cada base de datos trata o en lo que se especializa, que es algún tema en particular, a menos que sea una base de datos de temas generales; hoy en día se tiene la posibilidad de buscar en bancos de información especiales para un solo tema o área, esto hace que el proceso de búsqueda y recuperación de información se vuelva más certero y rápido al buscar únicamente en el ámbito temático deseado, así se puede hablar de bases de datos:

- científicas
- tecnológicas
- sociales
- económicas
- jurídico-legales
- bio-médicas
- económico-financieras

1.1.5.2 POR EL ÁMBITO TERRITORIAL. Con esto me refiero a que cuando se tiene alguna base de datos que está limitada por alguna cuestión como cuando se necesita determinar un tema de acuerdo al lugar ya que muchas veces se tiene, por ejemplo, información de revistas y periódicos nacionales pero se requiere buscar un tema de otro país, entonces se remite a una base de datos internacional o bien, del país mismo, por ello se hizo la siguiente clasificación de las bases de datos desde el punto de vista territorial:

- internacionales
- nacionales
- regionales
- locales

1.1.5.3 POR SU ACCESO. En ocasiones se tienen bases de datos a las que no es posible acceder pues contienen información privada para lo cual se necesita un password/contraseña o tener un convenio y/o suscripción con la empresa que la administra; un ejemplo típico sería la información financiera, información secreta de planes, procedimientos y estrategias de una empresa particular. También se encuentran bases de datos de dominio público en la que toda la gente puede tener acceso, ejemplo de ello sería una red de redes, INTERNET, en donde algunas páginas son privadas pero la mayoría de ellas son de dominio público y sin restricción. Por lo tanto, puede decirse que existen:

- bases de datos privadas o restringidas.
- bases de datos de acceso abierto o públicas.

1.1.5.4 POR SU ORIGEN O PROCEDENCIA. En muchos casos se trata de bases de datos hechas por alguna institución o asociación determinada, lo que permite saber qué tipo de información contiene. Estas pueden dividirse en:

- institucionales
 - públicas
 - privadas

1.2 MODOS DE ACCESO A LAS BASES DE DATOS

Para finalizar estas referencias al concepto y tipos de bases de datos, es conveniente definir los diferentes modos de acceso a las mismas:

1.2.1 CONEXIÓN "EN LÍNEA"

La denominación teledocumentación corresponde al conjunto de técnicas provenientes de la documentación científica, la informática y las telecomunicaciones que permiten el acceso y recuperación a distancia de la información.⁷

En cuanto a la línea de telecomunicaciones; "Dícese de todo medio físico, por ejemplo un hilo conductor o un haz de microondas, que se emplea para transmitir datos. Sinónimo de línea de transmisión."⁸

Para hacer más clara la definición, debe decirse que el término "telecomunicaciones": se relaciona con la emisión o recepción de señales, imágenes, sonidos o información por soportes de hilos metálicos, radio, fibra óptica o enlaces de infrarrojo. Se deriva del griego, donde el prefijo "tele" significa "lejos".⁹

⁷ Coll-Vinent, Roberto. Banco de datos: teoría de la teledocumentación. Madrid : Ate, 1980. p.17.

⁸ Held, Gilbert. Diccionario de tecnología de las comunicaciones. Madrid: Wiley, 1997. p.304.

⁹ Ibidem p.550

En el proceso de la conexión en línea, el esquema simplificado es el de una computadora central en el que se hallan almacenadas las bases de datos que son accesibles mediante terminales remotas.

Habitualmente estas terminales son calificadas con el término "dedicadas" en el sentido de que se emplean simplemente como periféricos de entrada-salida conectados a una computadora central; últimamente, la tendencia es utilizar para esta labor computadoras personales con diversos programas que facilitan el acceso a las bases de datos.

Abordando la cuestión con un poco más de detalle, habitualmente se habla de que son cuatro los elementos de la cadena de teledocumentación: productores de bases de datos, distribuidores, redes de transporte y agentes o mediadores de la información.¹⁰

- Los primeros están constituidos por aquellas empresas o instituciones que deciden informatizar sus archivos documentales, transformándolos en bases de datos y poniendo las mismas a disposición de la comunidad de usuarios.
- El segundo eslabón de esta cadena está formado por los distribuidores (Host en la terminología inglesa y serveurs en la francesa). Se trata de organizaciones que almacenan la información en sus sistemas informáticos y la hacen recuperable mediante la conexión de terminales a los mismos. Tal conexión es posible gracias a la existencia de las redes de telecomunicación normalmente pública, al menos en Europa.
- Estas redes posibilitan la transmisión de información, bien mediante sistemas especificados de comunicación de datos, bien a través de redes de uso general

(por ejemplo, telex o la red telefónica conmutada).

- El último elemento, el más cercano al usuario, lo constituyen los agentes o mediadores de la información. Esta actividad es desarrollada habitualmente en los servicios de información y documentación de empresas e instituciones, y en ellos personal especializado realiza la conexión y consulta a las diversas bases de datos en función de las demandas formuladas por los usuarios. Todo este complejo entramado constituye un auténtico sector cuya importancia no ha dejado de crecer.

Algunos datos permiten ofrecer una idea más exacta de la estructura de este sector.

Con respecto a la temática o contenido de las bases de datos, la mayoría de las mismas, casi el 40%, pertenecen al ámbito económico-financiero, le siguen con el 35% las de tipo científico y técnico, y, por último en 15%, las centradas en el área de las ciencias humanas y sociales.

Al indagar el origen institucional de las bases de datos, es posible observar cómo, mientras en Europa la mayoría de las bases de datos son públicas, en Estados Unidos y Canadá la producción y distribución de las mismas depende de empresas privadas.

1.2.2 ACCESO LOCAL

La posibilidad de acceder "fuera de línea" (*off-line*) a las bases de datos está dada por la capacidad de almacenamiento de los soportes ópticos y por el incremento de potencia y velocidad de proceso de las computadoras personales. CD-ROM es un acrónimo cuyo desarrollo en inglés es *Compact Disc- Read only Memory* y que puede traducirse como "Disco Compacto- Memoria únicamente de lectura". *Read Only*

Memory es un término genérico usado para designar cualquier dispositivo informático de almacenamiento de datos permanente y no modificable.¹¹

A veces se hace excesivo hincapié en la enorme capacidad del CD- ROM olvidando que éste es un circuito periférico más de la computadora, y que es este último el que debe localizar y procesar tal masa de información.

Al margen de que el desarrollo tecnológico lo haya posibilitado, al auge del acceso local han contribuido, además, otros factores primordialmente de índole económica.

En primer lugar, el incremento de los costos de conexión a las bases, tanto en lo relativo a las tarifas de los distribuidores como a las de los servicios de telecomunicación. Aunque actualmente el costo de las bases de datos en CD-ROM no es especialmente económico, es previsible que tienda a la baja. En este momento, la adquisición o suscripción a las mismas resulta costeaible para aquellas instituciones que realicen consultas en línea frecuentemente.¹²

Por otro lado, el acceso local permite a las editoriales o productores-distribuidores un más eficaz control de la política de precios, y al tratarse de un soporte de sólo lectura, se impide el acceso no permitido.

Sumado a esto, y al margen de las consideraciones económicas, el acceso local permite al usuario no estar pendiente del tiempo de conexión y no depender de la calidad del servicio telefónico como las búsquedas en Internet.¹³ Este último punto es

11 Held, Gilbert. Op. Cit. p.432.

12 Hlava, Marjorie M.K. "CD-ROM vs. Online: can both exist?" p: 14. En: Bulletin of the American Society for Information Science. Vol. 14, no. 1 Oct.-Nov., 1987. P. 45-53.

13 Mahalingam, K.; Huhns, M.N. "A tool for organizing Web information". En: Computer, June, 1997. Vol. 30, p: 80 -83.

de especial importancia para los usuarios de algunos países donde el servicio que prestan las compañías telefónicas no es precisamente ejemplar. Desde el punto de vista de los editores de CD-ROM tales países deben ser vistos como auténticos paraísos llenos de prometedoras posibilidades comerciales.

Actualmente el éxito de las computadoras personales se debe, entre otras muchas razones, a que son los propios usuarios los que, sin la mediación de los técnicos de los centros de cálculo, realizan sus propias aplicaciones.

Este modo de operar ha llevado a un esfuerzo cada vez mayor por parte de los diseñadores de lógicas para elaborar productos cada vez más "amigables", más "intuitivos" y sencillos de manejar. Esto mismo ha ocurrido en el sector de las bases de datos. La mayoría de los programas de recuperación de la información que incorporan las bases de datos en CD-ROM son de manejo más sencillo que sus equivalentes de los sistemas en línea.¹⁴

Por otro lado, no hay que olvidar que por muy amigable y bien diseñado que esté, por ejemplo, un paquete estadístico (como el Statistical Package for Social Sciences) el usuario del mismo debe saber estadística para usarlo. De hecho, siempre se dice que un programa informático debe elaborarse de forma que el usuario no tenga que preocuparse de las cuestiones informáticas para centrarse en lo específico del campo de aplicación del mismo.

El acceso en disco permite no aprender a conectarse a las bases, pero, como ya quedó dicho, por muy bueno que sea el programa de recuperación, no exime al usuario de tener algunos conocimientos de documentación automatizada para

¹⁴ Sherman, Chris. The cd rom: Handbook. New york: Intertext, 1988. p:37.

optimizar los recursos disponibles.

En cualquier caso, y en el ámbito específico de las bases de datos, lo que parece claro es que se trata de dos modos de acceder a la información, que si bien a primera vista parecen antagónicos deben tender, y de hecho así está ocurriendo, a complementarse. Esta complementación se justifica si se consideran algunos elementos tales como el tamaño de la base, la frecuencia deseable de actualización y, desde el punto de vista del usuario, la frecuencia de las consultas a la misma.

Parece obvio que, independientemente de los precios vigentes y de las "facilidades" telefónicas, no merece la pena adquirir una base de datos en CD-ROM que se consulta de vez en cuando. Los ciclos más habituales de las bases en CD-ROM son el semestral y el trimestral. La actualización se lleva a cabo mediante el envío de un disco nuevo y la correspondiente devolución del "antiguo".

Para ciclos menores parece más apropiada la conexión en línea, aunque existen compañías que -para evitar un continuo trasiego de discos compactos- optan por realizar actualizaciones frecuentes en disquetes o unidades de 3½ o en discos ópticos una o dos veces al año.¹⁵

A la hora de efectuar las búsquedas, el sistema consulta, tanto al disco compacto como al disco duro en el que se han instalado las últimas actualizaciones. En cualquier caso, para información cuya utilidad sea vigente por periodos inferiores a una semana, parece claro que los servicios en línea son los idóneos.

Por último, queda por examinar la cuestión del tamaño de la base de datos. El soporte CD-ROM parece especialmente adecuado para bases de datos cuyo tamaño oscile

¹⁵ Desmarais, Norman. *The librarian's CD-ROM handbook*. Westport, Connecticut : Meckler, 1989, p.45.

entre los 20 y los 500 Mb. Bases de datos más pequeñas pueden ser instaladas sin problemas en un disco duro de capacidad creciente. A algunos quizá les extrañe comenzar el intervalo con un tamaño mínimo de 20 Mb. Evidentemente, si sólo tiene este tamaño y su tasa de crecimiento es baja para una base de datos de estas características resulta demasiado grande un disco compacto. Como en el mercado abundan las bases de datos de tamaño medio, lo que suelen hacer las empresas es distribuir varias de temáticas similares en un solo disco.

Con respecto a las grandes bases de datos, el límite viene impuesto por el del tamaño del CD-ROM. Como son también bastantes las que exceden de 500 Mb., lo que se acostumbra hacer, es distribuir las en varios discos. De esta forma, las empresas venden el "fondo histórico" de una base de datos en tantos discos como sea necesario, y la información más reciente con una antigüedad de dos o tres años, se saca al mercado como suscripción en el disco, que se actualiza periódicamente. Este es el modo de operar habitualmente a la hora de "trocear" una base cuyo tamaño excede lo que cabe en un solo disco y es que tiene más éxito en aquellos ámbitos en que la información "envejece" rápidamente.

El CD-ROM ha supuesto un cambio importante en el sector de las bases de datos. Hasta su aparición la consulta que el usuario hacía, con mayor o menor frecuencia, en forma premeditada, a través de mediadores y siempre condicionado por los tiempos de conexión.

Aquel usuario que tenga acceso a una computadora personal equipada con este

sistema, puede ahora realizar consultas tantas veces como desee, de forma autónoma y sin tener presente el reloj. Pero en general, al margen del número de personas que tengan que compartir la PC, su frecuencia de uso se incrementa y la consulta a las bases pasa a ser un recurso de trabajo habitual.

CAPÍTULO 2

RECUPERACIÓN DE INFORMACIÓN EN BASES DE DATOS

2.1 LA INFORMACIÓN Y SU IMPORTANCIA

El término información tiene su origen en las palabras in y formare, que significan, "instruir hacia adentro". A partir de estas dos palabras, y debido a la importancia que en épocas recientes ha cobrado el término "información", se ha generado una enorme cantidad de variantes, cada una con un significado muy preciso, aplicable a determinadas situaciones; por ejemplo, "informática"; apócope de "información automática" supone el procesamiento automático de la información; "telecomunicaciones" significa comunicar a distancia; "telemática" es la conjunción de "telecomunicaciones" e "informática", e implica la transmisión y el procesamiento automático de la información.¹⁶

Otra definición de información es la siguiente: "Significado atribuido a los datos por las personas. Contenido organizativo de una señal".¹⁷

"La información es coleccionable, almacenable o (sic) reproducible. Se utiliza para tomar decisiones, conduce también a conclusiones acertadas o equivocadas, puesto que puede ser interpretada de diversas formas por distintos individuos, dependiendo de muchos factores subjetivos y del contexto en que se encuentre la persona que la

¹⁶ Freedman, Alan. Diccionario de computación. Madrid: McGraw-Hill, 1993. p: 734.

¹⁷ Held, Gilbert. Op. Cit. p. 269.

recibe e interpreta. Así como es posible comunicar una noticia, también se comunican los estados de ánimo, opiniones o conocimientos. La información es la representación del conocimiento, tiene una naturaleza física, es producto y base del conocimiento, y representa una de las formas de comunicar la estructura del conocimiento a través de datos, la información es la forma tangible y comunicable del conocimiento.¹⁸

Actualmente es difícil pensar que alguien niegue conscientemente que la información tiene un valor fundamental; la información ha ido ganando importancia conforme las personas o líderes que toman decisiones están convencidos de que aquella se puede asociar a un valor real, frecuentemente ligado a un valor material o económico. En el final del siglo XX y principios del XXI es de tal importancia poseer, administrar y transmitir información, que toda la humanidad se ve y se seguirá viendo afectada, influida y posiblemente dominada por quienes tienen, administran y transmiten este recurso, razón por la cual a esta época se le han impuesto los calificativos de "sociedad de la información" o de "revolución electrónica", éste último debido a la facilidad con que se transmite la información por medio de los sistemas modernos basados en dispositivos electrónicos.¹⁹

Uno de los aspectos más abstractos e importantes de la información es su temporalidad porque valor puede disminuir a lo largo del tiempo. Es decir, en un momento determinado a alguien le puede interesar contar con cierta información, pero ese interés puede decrecer o incluso desaparecer algún tiempo después. Por otra parte, es necesario que la información sea de interés para el individuo que la

18 Kuhlmann, Federico. Información y telecomunicaciones. México: Fondo de Cultura Económica, 1997. p: 13.

19 Ibidem pp.78-79.

adquiera o reciba.

La información se origina en una fuente y se hace llegar a su destinatario por medio de un mensaje a través de un canal de comunicación; el destinatario generalmente se encuentra en un punto geográfico distante o, por lo menos, separado de la fuente. La distancia entre fuente y destinatario puede variar desde pocos centímetros (al hablar frente a frente a un volumen normal) hasta cientos y aún miles de kilómetros (como es el caso de transmisiones telefónicas intercontinentales o de transmisiones desde y hacia naves espaciales).

Shannon,²⁰ quien es un experto de este tema describe con sencillez asombrosa el proceso de la información y menciona que un sistema de comunicaciones consiste de cinco componentes:

- 1) una fuente de información;
- 2) un transmisor de información cuya función consiste en depositar la información proveniente de la fuente en un canal de comunicaciones;
- 3) un canal de comunicaciones, a través del cual se hace llegar la información de la fuente a su destino;
- 4) un receptor que realiza las funciones inversas del transmisor, es decir, extrae la información del canal y la entrega al destinatario;
- 5) un destinatario.

²⁰Shannon, Claude E. *The mathematical theory of communication*. Urbana: University of Illinois, 1969. p.65.

El mensaje hace llegar la información desde la fuente hasta su destino, y no es lo mismo un mensaje que la información que contiene. Considérese el siguiente ejemplo: Una persona (A) desea enviar cierta cotización para el cierre de un negocio por medio de una carta a otra persona (B). En este caso, A es la fuente, B el destinatario. La información es aquello necesario para conocer la cotización y para originar la entrega del mismo a B, y el mensaje es el conjunto de palabras escritas en la carta de la cotización necesarios para que B conozca la intención de A y para que B pueda conocer la propuesta de el negocio que A le envía.

Desde los orígenes de la humanidad, la forma natural en que la información se transmite entre las personas es a través del lenguaje oral; además, existe una etapa de transmisión escrita. En la actualidad, también existe la necesidad de transmitir información a través de computadoras, ya que por cuestiones de tiempo, distancia y exceso de trabajo, el ser humano se ve obligado a utilizar las nuevas tecnologías de información y así optimizar los recursos disponibles; por lo tanto hoy en día no es necesario intercambiar información oralmente ya que existe la infraestructura y los medios para llevar acabo el proceso del intercambio de la información de manera más fácil y adaptable a este fin de siglo.

2.2 LA RECUPERACIÓN DE LA INFORMACIÓN

Con esta expresión (que en la literatura anglosajona aparece como "information retrieval") se designan todas las técnicas relativas al acceso y posterior selección de la

información almacenada en un sistema²¹.

Todo el trabajo que hacemos los bibliotecarios y los especialistas en el campo de la información tiene como fin que la información que se procese e ingrese en la máquina y/o en la red sea utilizada, sea empleada por algún usuario alguna vez; existe información que nunca es llamada a una pantalla, y que nunca es utilizada sería un desperdicio dado el alto costo de su producción, organización y de almacenamiento.

La recuperación de la información es un proceso que realiza el usuario con las bases de datos, el cual está orientado a identificar aquellos registros o documentos relevantes para el usuario.

La actividad de recuperación de la información, no siempre logra realizarse. Si se realiza, no siempre se alcanza el éxito deseado, aún en el caso de que la información haya sido ingresada en una fuente o sistema, se haya solicitado, y se haya intentado su búsqueda en repetidas ocasiones y por las diferentes vías al alcance de quien consulta, si es que no se tienen estrategias y técnicas bien definidas que le ayuden a la búsqueda de información.

Hasta la fecha, todos los servicios de recuperación usan metodologías conceptualmente similares y muchas veces son apoyadas desde las mismas bases de datos, ya que mantienen la "ayuda" activada para que el usuario pueda aclarar sus dudas. Desde el punto de vista bibliotecológico existen muchos instrumentos que habilitan a los usuarios de bases de datos para recuperar información. Estos instrumentos son las llamadas "guías o manuales del usuario", que manejan

²¹ Lizasoain, Luis. Bases de datos en CD-ROM. Madrid : Paraninfo, 1992. p.109.

descriptores y temas. En el área de cómputo, esta actividad ha estado basada en el álgebra binaria, la lógica booleana y otras técnicas muy útiles en el ámbito informativo que más adelante se abordarán.

2.3 RECUPERACIÓN DE LA INFORMACIÓN EN BASES DE DATOS

Cuando un usuario decide utilizar este tipo de recursos, lo primero que debe hacer es formular la pregunta con su propia terminología.²²

Esta formulación debe incluir conceptos o ideas básicas, así como las relaciones que entre aquellos se establecen, expresados en lenguaje natural.

La segunda tarea a realizar es, consecuentemente con lo anterior, seleccionar la o las bases de datos en las que previsiblemente se encuentra almacenada la información relativa al tema que se investiga.

Dicha selección suele efectuarse tomando en consideración criterios tales como el tipo de documentos almacenados, la procedencia de los mismos, la cobertura temporal, etc.

Aunque en determinadas áreas existen bases de datos que son casi de referencia obligada habida cuenta de lo amplio de su cobertura espacio-temporal y de la calidad del proceso de selección de los documentos, existen otras bases de datos cuya consulta a veces no resulta de utilidad por lo excesivamente especializado de su campo de actuación o por lo restringido de la procedencia de las fuentes.²³ Aunque para algunos usuarios si puedan ser útiles, de lo contrario no se justificaría su

22 Echeverría, L.E. "An intuitive approach for the expression of Boolean queries". En: Visual Languages. IEEE 1989. p: 118 -123

23 Frants, Valery I. "Boolean search: Current state and perspectives". Journal of the American Society for Information Science. New York, Jan 1999. p: 55.

existencia.

Una vez realizada esta selección, sólo queda recuperar la información de forma que los resultados de este proceso proporcionen un conjunto de referencias lo más ajustado posible a la demanda inicialmente formulada.

Pero la eficacia del sistema no depende sólo de que la información recuperada sea considerada como pertinente, es necesario además, que sea exhaustiva. En consecuencia, se trata de recuperar toda la información relativa al asunto solicitado, y sólo ésa. Evidentemente, como luego se verá, el obtener a la vez el máximo de relevancia y el máximo de exhaustividad, es poco menos que imposible.

Como ya se ha comentado, los productores y administradores de las bases de datos han dotado a las mismas de una serie de recursos informáticos y documentales que - si son correctamente empleados- facilitan la labor. Una vez finalizado el proceso de selección de las bases de datos que se suponen apropiadas a la demanda de información formulada, la siguiente tarea, previa a la realización de la consulta, está centrada en el conocimiento de dichas bases, de su estructura, contenido y procedimiento o procedimientos de indización que adopta.

En definitiva, se trata de saber cuál es el modelo de los registros, nivel de estructuración, cuántos y de qué tipo son los campos que contiene y cómo se puede recuperar la información en ellos contenida.

Al mismo tiempo, es preciso conocer las características, posibilidades y modos de programación y operación del lenguaje informático de recuperación. Esto generalmente es diferente para cada base de datos aunque es posible encontrar

algunas que a pesar de esa diferencia comparten el mismo sistema informático.

Además, hay que tener presente que la mayoría de estos lenguajes tienen una estructura y funcionamiento similar. Por último, si como es habitual se trata de una base de datos bibliográfica, lo normal es que ésta cuente en su estructura con campos de índices que pueden adoptar diversas formas: códigos, alfabéticos o alfanuméricos, de clasificación temática, palabras-clave, frases-clave, descriptores, etc. Estos son campos que están específicamente diseñados por el productor de la fase para recuperar la información contenida en la misma.²⁴

El usuario debe conocer los procedimientos de indización usados, con el fin de adaptar su demanda de información a las reglas del lenguaje documental empleado en muchas de las bases de datos bibliográficas.

Se trata, en definitiva, de que el usuario que intenta recuperar información sobre un determinado tema emplee en dicho proceso los mismos términos, o códigos numéricos que han sido empleados al registrar en la base los documentos relativos a esa cuestión.

Si por ejemplo, en una base de datos sobre química utilizan un código numérico de clasificación de forma que a toda la documentación relativa a los polímeros corresponda el código 872.21, el usuario interesado en recuperar documentos que traten de forma genérica esta cuestión debe conocer dicho código.

Análogamente, si en una base de datos han optado por emplear el término "ordenador" en vez de "computadora", si se intenta recuperar documentos indizados con esta última palabra, la consulta está destinada al fracaso.

²⁴ Bodoff, David. "A re-unification of two competing models for document retrieval". *Journal of the American Society for Information Science*; New York; Jan 1999.

En conclusión, se puede decir que, una vez seleccionadas las bases de datos que se consideran adecuadas, el proceso de recuperación de la información pasa por conocer con detalle la estructura de cada base y, sobre todo, consiste en traducir la demanda de información inicialmente formulada en el lenguaje natural del usuario a una ecuación de búsqueda expresada en un lenguaje formal.²⁵ Esta labor de "traducción" lo es en un doble sentido: en primer lugar, hay que considerar las reglas, recursos y procedimientos del lenguaje informático que gestiona la recuperación de la información en cada base de datos. Y, en segundo, los conceptos inicialmente incluidos en la demanda deben ser traducidos a los términos propios del lenguaje documental empleado al indizar la información de cada sistema, de acuerdo también a las reglas propias de este tipo de lenguajes.²⁶

Desde esta doble perspectiva se establece el siguiente paralelismo:

La demanda de información formulada por el usuario está formada por los conceptos y las relaciones que existen entre ellos expresados en lenguaje natural.²⁷

Por ejemplo: documentos sobre el empleo de computadoras en los procesos de control de calidad que estén escritos en inglés con posterioridad al año 1988.

La ecuación de búsqueda que se introduce al sistema para recuperar información consta de términos de indización y operadores y está expresada en lenguaje formal.

Por ejemplo: <COMPUTERS AND QUALITY-CONTROL) IN DESCRIPTORS) AND LANGUAGE = ENG AND YEAR > 1988.

²⁵ Bordogna, G. "Extending Boolean information retrieval: a fuzzy model based on linguistic variables". *Fuzzy Systems, 1992. IEEE International Conference, 1992.* p: 769 -776.

²⁶ Keim, Ichelle Renee. "Bayesian information retrieval (database, text retrieval, meta-analysis)". University of Washington. PDH, 1997. p. 56

²⁷ *Ibidem* p. 769.

El siguiente esquema sintetiza lo dicho:

DEMANDA DE INFORMACIÓN		ECUACIÓN DE BÚSQUEDA	
Conceptos	(Lenguaje natural)	Términos	(Lenguaje documental)
Relaciones	(Lenguaje natural)	Operadores	(Lenguaje informático)

Como se puede observar se trata de encontrar la manera adecuada de "traducir" el lenguaje natural de la búsqueda de información que el usuario realiza al lenguaje documental que la base de datos emplea (términos y operadores que se relacionen). Insisto en lo de la simplificación porque este doble proceso de traducción no implica procedimientos separados entre sí, entre lo relativo a los lenguajes informáticos y a los documentales. Como enseguida se verá, ambos campos comparten principios y procedimientos.

Según lo dicho en este capítulo se verá más adelante en un ejemplo ficticio la estructura típica de una base de datos bibliográfica y los métodos y técnicas fundamentales de la recuperación automatizada de la información.

2.3.1 TIPOS DE RECUPERACIÓN DE LA INFORMACIÓN EN LAS BASES DE DATOS

El primer capítulo ya se había dado una introducción a las bases de datos; se examinaron a fondo lo que son las bases de datos factuales o numéricas y las bases de datos textuales o de recuperación documental, entonces es oportuno hablar sobre "recuperación factual" y "recuperación documental".

En el primer caso, se trata de buscar una información concreta habitualmente de tipo numérico. Por ejemplo: volumen de la producción de trigo en México durante el año

1990. Se trata de consultas de tipo determinista que habitualmente se efectúan sobre bases de datos numéricos.

Por el contrario, la recuperación documental es de tipo probabilístico y se lleva a cabo normalmente sobre bases de datos documentales o textuales.

- En el primer caso, la respuesta del sistema es una cantidad.
- Por el contrario, en el segundo, el sistema proporcionaría una lista de documentos. Puede que en algunos de dichos documentos se encuentre el dato requerido en la consulta del primer tipo, es decir, de recuperación factual, o que simplemente se analice la problemática del sector o los fertilizantes más empleados en la producción de trigo en el año de 1990.

En lo antes dicho se encuentra la razón de la distinción establecida entre recuperación factual y documental. En el primer caso, es posible valorar inmediatamente por parte del usuario la exactitud de la respuesta y, por ende, la eficacia del sistema.

Por el contrario, en un sistema documental más que de exactitud se habla de utilidad, de pertinencia en función de las necesidades de información del usuario. La recuperación de un documento en una base de datos es un medio para localizar un contenido particular representado por una parte del conjunto de documentos que conforman la base de datos. El contenido de la base de datos es el **qué** de la recuperación de información, el subconjunto en el que tal contenido reside es el **dónde**.

- La consulta es una operación que va del dónde buscar la información al qué tipo de información se está buscando el usuario se introduce a la base de datos en un punto, en un pasaje de la misma, y mediante desplazamientos por el conjunto de la

información almacenada va localizando los registros y selecciona el que le interesa o la que se adapta a sus necesidades de información.

- La interrogación toma el sentido contrario, va del qué al dónde; el usuario con los términos de búsqueda y mediante ellos localiza el subconjunto de documentos que se ajustan a las necesidades de información personales.

La consulta es el modo habitual al usar las denominadas precisamente, obras de consulta. El usuario abre el texto por el índice, y en función de los resultados que obtuvo se va desplazando a los diferentes pasajes o capítulos que le interesan y así puede elegir los registros más significativos. Por extensión y analogía, es el modo de proceder habitual en las bases de datos en texto completo y en general en los productos típicos de edición electrónica tales como enciclopedias, obras de referencia, etc.

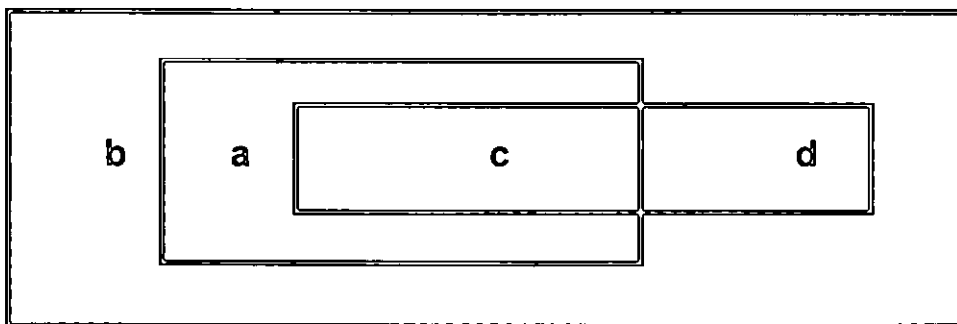
Para seguir con el símil, la interrogación es como el proceso que sigue un usuario cuando abre una enciclopedia por el índice temático o alfabético de materia, sabe qué palabras o conceptos busca y en dicha tabla encuentra las páginas en que aparecen. Este es el modo habitual de operación con las bases de datos estructuradas de tipo bibliográfico. En primer lugar, se interroga al sistema mediante palabras-clave o descriptores y, como resultado, se obtiene un primer conjunto de documentos. Después de la lectura de los registros de los mismos se puede volver a consultar la base con el fin de ampliar o restringir la búsqueda en función de determinados objetivos y empleando los índices inversos formados por todos los términos del texto de los documentos. Se suele hablar, entonces de consulta y de interrogación a las bases de datos.

2.3.2 INDICADORES DE LA EFICACIA DE LA RECUPERACIÓN

Independientemente, de que sea una cuestión subjetiva la evaluación que el usuario realiza, de lo adecuado o no a sus necesidades de información, el conjunto de documentos recuperados, cuando se habla de eficacia en este contexto es preciso abordar dos cuestiones: **la precisión y la exhaustividad** o, dicho en otros términos, **ruido y silencio**:

El ruido está formado por las referencias no útiles obtenidas en una búsqueda. La tasa de ruido es el porcentaje de referencias no útiles sobre el total de las obtenidas. A la noción de ruido se opone la de silencio, que se define como el número de documentos pertinentes existentes en el fondo documental y que no han sido seleccionados en la búsqueda. La tasa de precisión en la búsqueda es inversa a la del ruido y surge de la proporción de referencias pertinentes sobre el conjunto recuperado. La tasa de exhaustividad es el porcentaje de las referencias pertinentes obtenidas en la consulta sobre el total de los pertinentes existentes en el fondo documental.

El conjunto de referencias que constituye el fondo documental de un sistema lo se puede representar según el siguiente modelo que propone Lizasoani:



a = documentos pertinentes recuperados

b = documentos pertinentes no recuperados (silencio)

c = documentos no pertinentes recuperados (ruido)

d = documentos no pertinentes no recuperados

a + b = documentos pertinentes existentes

a + c = documentos recuperados

a + b + c + d = total de documentos en el fondo

$$\text{Tasa de exhaustividad} = \frac{a}{a + b}$$

$$\text{Tasa de silencio} = \frac{b}{a + b}$$

$$\text{Tasa de pertinencia} = \frac{a}{a + c}$$

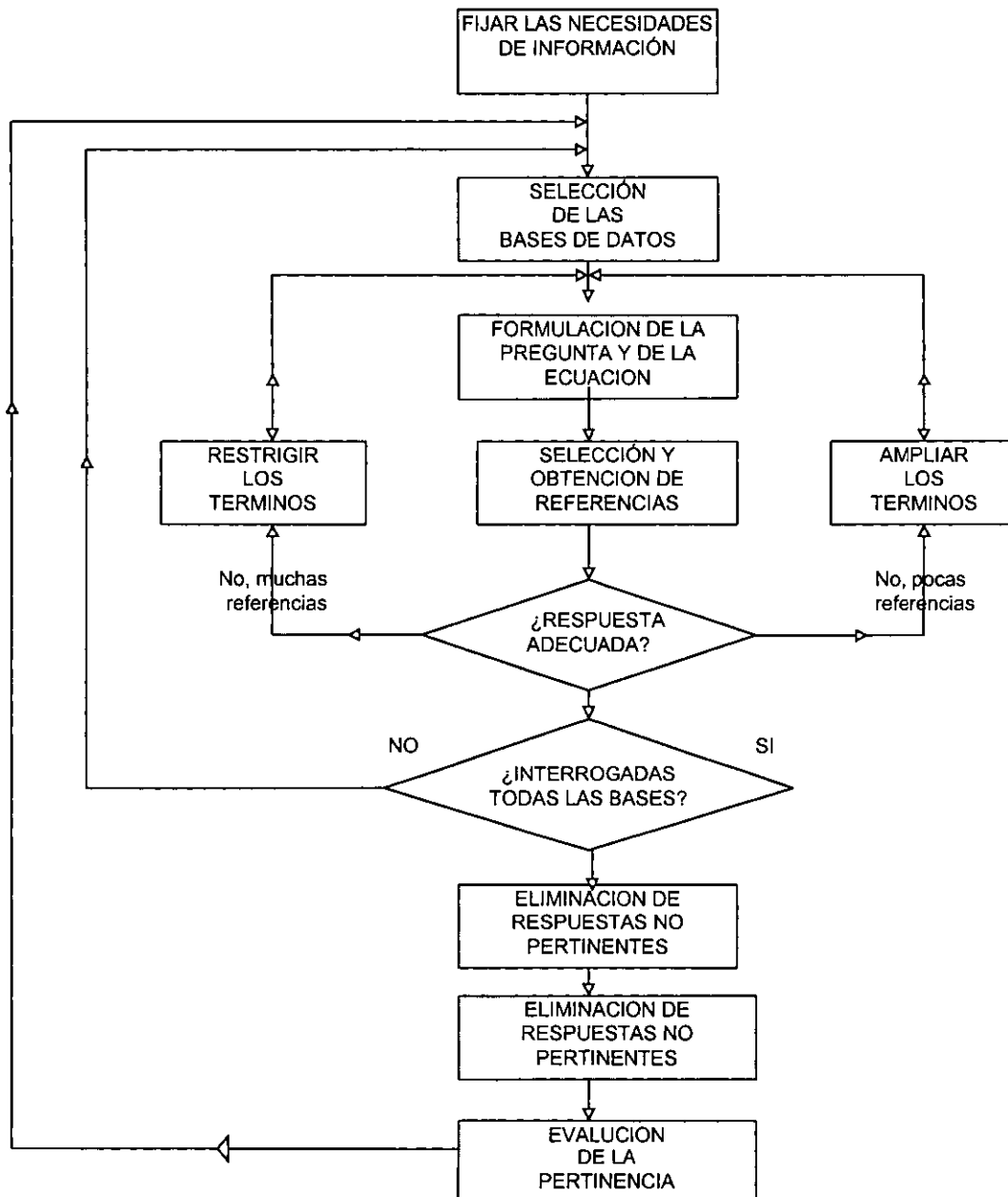
$$\text{Tasa de ruido} = \frac{c}{a + c}$$

Evidentemente, el objetivo ideal es proporcionar al usuario todos los documentos pertinentes y sólo éstos (ausencia de ruido y de silencio).

Como ya se dijo, para optimizar en lo posible los resultados de la recuperación de información, deben emplearse correctamente todos los recursos, informáticos y documentales con los que cuentan estos sistemas.

2.4 MODELO DE MÉTODOS Y TÉCNICAS DE RECUPERACIÓN DE INFORMACIÓN

En este apartado se expone a través de un diagrama propuesto por Lizasoani el modelo de los métodos y técnicas de recuperación de información en general, de manera ilustrativa y explícita se puede analizar paso por paso el proceso de búsqueda de información con base en el propio diagrama guiar el proceso de búsqueda.



2.5 LENGUAJES INFORMÁTICOS, MÉTODOS Y TÉCNICAS

El procedimiento más simple para llevar a cabo la interrogación al sistema consiste en teclear un término junto con la instrucción adecuada del lenguaje informático que se emplee, por ejemplo:

FIND COMPUTERS,

Pero tal como se ha venido exponiendo, la mayoría de las bases de datos y de los lenguajes de recuperación actualmente permiten efectuar consultas más sofisticadas y cuentan con herramientas y aplicaciones útiles para mejorar la eficacia en la recuperación de información.

Independientemente del tipo de lenguaje documental que se emplee, la interrogación a la base se puede efectuar de un modo más preciso si en vez de realizar este tipo de consultas simples, la interrogación se hace más compleja. En definitiva si se formula una ecuación de búsqueda como la siguiente:

ECUACIÓN = TERMINOS + OPERADORES

Los términos de búsqueda pueden ser "palabra clave" o descriptores, en función del tipo de lenguaje documental que se emplee, sea éste libre o controlado o tipo thesaurus.

La consulta a una base de datos, consiste en seleccionar del fondo documental existente en el sistema un subconjunto de referencias que reúnan determinadas condiciones requeridas por el usuario.

Estas condiciones (criterios de selección) se centran fundamentalmente en el contenido conceptual de los documentos, pero también suele hacer referencia a

aspectos tales como: la fecha de publicación de los mismos, la lengua en que están escritos, la revista o editorial donde se han publicado, etcétera.

Las relaciones y condiciones que ligan a los términos se expresan mediante los operadores, pero hay que tener una visión clara de los procedimientos y lenguajes documentales adoptados por los diseñadores de cada base de datos; en lo tocante a los operadores, estos se emplean en función de los lenguajes informáticos que poseen los diferentes sistemas.

A continuación se hará una descripción y categorización de los sistemas más habituales

2.5.1 OPERADORES BOOLEANOS O LÓGICOS

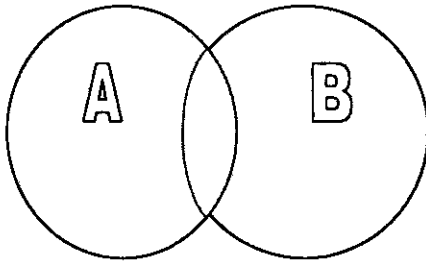
El empleo de los recursos del álgebra booleana constituye, sin lugar a dudas, uno de los instrumentos más poderosos para mejorar la eficacia en la recuperación de información. Se trata de un tipo de operadores universalmente empleados, y que es posible contar con ellos en todas las bases de datos y en los distintos lenguajes de recuperación. Estos operadores le indican a la computadora la relación lógica existente entre los términos buscados al nivel del registro. Esta lógica permite realizar tres tipos de enlace descriptores que se encuentran en los términos de búsqueda. Las principales relaciones lógicas usadas son:

- a) Unión de conjuntos: puede ser disyunción, "o" inclusión** e indica que las palabras son sinónimos o términos relacionados. Al menos un término deberá estar presente en los registros recuperados. Utilizar el operador "O" permite enlazar dos descriptores, debiendo hallarse uno u otro, o ambos en la indización del documento para que éste sea pertinente. Es usado para

incluir sinónimos, para ampliar los términos de búsqueda y también para eliminar duplicación de registros.

El siguiente ejemplo aclara lo anterior:

La Intersección, $(A \wedge B)$ (A y B), permite obtener del fondo documental todos los documentos caracterizados por A o B.



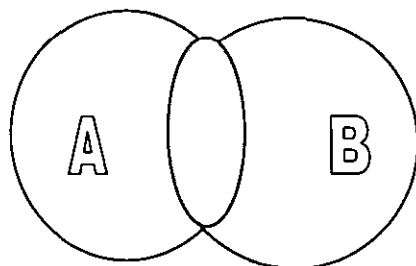
Ejemplo: todos los documentos que tratan de "Informática" (A) o "Medicina" (B).

La exclusión o negación $(A \vee B)$ permite obtener del fondo documental todos los documentos que en el análisis hayan sido caracterizados por A, salvo los que además lo hayan sido por B.

b) Intersección de conjuntos: conjunción, expresado por "y", se usa para combinar conceptos, en donde todos deberán estar presentes en el mismo registro pero no precisamente en el mismo campo. Utilizar el operador "Y" permite enlazar dos descriptores que deben encontrarse en la indización del mismo documento para que éste sea considerado pertinente. Concretamente, ambos términos deben encontrarse en cada registro. Este operador busca por artículos que contengan todas las palabras claves seleccionadas, lo cual resulta una limitante en la búsqueda.

Ejemplo: todos los documentos que traten de "Informática y de "Medicina" (B), o de "Informática y Medicina".

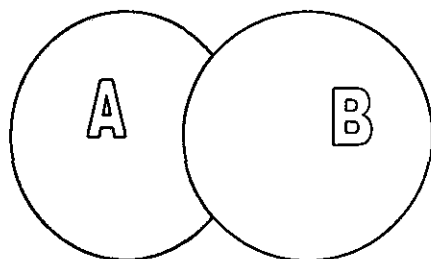
La Intersección, $(A \wedge B)$ (A y B), permite obtener del fondo documental todos los documentos caracterizados por A y B.



c) Exclusión de conjuntos: negación, expresado por "no" se utiliza para eliminar conceptos.

Cada conjunto está definido por una característica de análisis. El utilizar el operador "NO"; permite unir dos descriptores de los cuales el primero debe estar presente y el segundo ajustarse en la indización de un documento para que éste sea pertinente. Elimina doble significados de palabras y también limita la búsqueda.

La exclusión o negación $(A \vee B)$ permite obtener del fondo documental todos los documentos que en el análisis hayan sido caracterizados por A, salvo los que además lo hayan sido por B.



Ejemplo: todos los documentos que tratan de "Informática" excepto los que además traten de "Medicina", o que traten de "Informática" (A), salvo los que estén escritos en alemán (B).

Como resumen final, el siguiente cuadro reúne las tres diferentes denominaciones de operadores booleanos más utilizados y cómo se pueden llegar a expresar en los diferentes sistemas de recuperación de información:

INTERSECCIÓN, conjunción, producto: *, AND, Y, ET

UNIÓN, disyunción, suma: +, O, OR, OU

EXCLUSIÓN, negación, resta: NOT, NO, SAUF

2.5.2 OPERADORES DE EXPANSIÓN

Son aquellos que sirven para incrementar el campo de acción de la búsqueda a más de un término mediante la sustitución de caracteres en el mismo. Estos operadores de expansión suelen dividirse en dos: enmascaramiento y operadores de truncamiento. Los primeros sirven para reemplazar un solo carácter del término de búsqueda. Por ejemplo:

MICRO?COMPUTADORA

En este caso, el carácter "?" cumple la función de "comodín" de forma que equivale por cualquier carácter (incluida la secuencia vacía), siempre que sea sólo uno.

De esta manera se recuperarían simultáneamente los términos MICROCOMPUTADORA Y MICRO-COMPUTADORA.

Su empleo está especialmente indicado en las variaciones y formas alternativas de un mismo término. Los operadores de truncamiento se emplean habitualmente con las raíces de los términos. El símbolo más frecuente suele ser el asterisco (*). Por ejemplo, PROCES*. Este término nos serviría para recuperar palabras como PROCESAMIENTO, PROCESO, PROCESOS...

Este operador, por tanto, sustituye a cualquier secuencia de caracteres que le sigan.

2.5.3 OPERADORES DE INTERVALO

También se le llaman operadores relacionales, sirven para construir intervalos numéricos.

Los más habituales son:

IGUAL	=
DISTINTO	<>
MAYOR QUE	>
MENOR QUE	<
MAYOR O IGUAL	>=
MENOR O IGUAL	<=

Su uso es frecuente en el campo de la fecha o en otros campos de tipo numérico o temporal. Por ejemplo, AÑO >= 1989. En este caso, se obtendrían las referencias de los documentos publicados desde 1989, éste incluido, hasta la actualidad.

2.5.4 OPERADORES DE PROXIMIDAD O DISTANCIA

Se emplean, sobre todo, en los campos de texto libre o en las bases de datos en texto completo. Sirven para definir la distancia máxima admisible que debe separar a

los términos. La búsqueda de proximidad es muy útil cuando se está interesado en temas que son normalmente conexos pero no son representados necesariamente con una frase exacta del idioma.

Habitualmente adoptan la forma de un número que indica el máximo número de términos intermedios.

Por ejemplo: INFORMÁTICA D= 3 DOCUMENTACIÓN

Aquí se obtendrían todos aquellos registros en los que figuraran estos dos términos separados como máximo por tres términos o conceptos.

Por ejemplo esta frase: INFORMÁTICA APLICADA A LA DOCUMENTACIÓN

Pero no, LA DOCUMENTACIÓN ES UNO DE LOS CAMPOS DE APLICACIÓN DE LA INFORMÁTICA

Ya que en este caso los conceptos están interrelacionados por más de tres términos, porque después de "documentación" están las palabras "es uno de los campos de" y hasta después está "aplicación" e "informática", por lo tanto no se cumple la condición.

2.5.5 OPERADORES DE CUALIFICACIÓN

Aquí se trata de la posibilidad que brindan la mayoría de los lenguajes de recuperación de especificar el campo en el que debe verificarse la condición, expresado mediante el empleo del resto de los operadores.

Hasta ahora, y dado que no se había especificado lo contrario, las búsquedas serían realizadas sobre todo el texto de los registros de la base correspondiente.

Así, el ejemplo de la intersección (INFORMÁTICA Y DOCUMENTACIÓN), daría como resultado que se obtuviéseren como respuesta las referencias de todos los documentos en los que aparezcan simultáneamente estos dos términos, pero independientemente del lugar de registro en que se encuentren.

Así, podría haberse dado el caso de que "INFORMÁTICA", aparéciere en el título y "DOCUMENTACIÓN" en el resumen. Si se quiere especificar más y obtener una respuesta más precisa, se podría cualificar la búsqueda poniendo como condición que estos dos términos aparezcan en el título o en los descriptores o temas:

(INFORMATIC AND DOCUMENTATION) IN TITLE,

(INFORMATIC AND DOCUMENTATION) IN DESCRIPTORS.

De esta manera se formula una expresión más compleja mediante el empleo simultáneo de varios operadores. Esta posibilidad de definir el campo de búsqueda aumenta la precisión de las búsquedas y además disminuye el tiempo de respuesta. De tal forma la búsqueda se lleva a cabo sobre este índice específico y no sobre el conjunto del texto del registro.

2.5.6 OPERADORES DE ENCADENAMIENTO

Por último, la mayoría de los lenguajes de recuperación permiten emplear como elementos de una ecuación de búsqueda los resultados de las búsquedas anteriores.

Así se pueden ir efectuando búsquedas simples y en función de los resultados de las mismas, combinarlos para obtener una respuesta final satisfactoria.

Por ejemplo, si se desea obtener información sobre documentos que traten acerca de: "Aplicaciones del CD-ROM en bases de datos que sean posteriores a 1989 y que no estén escritos en alemán".

Una posibilidad es realizar una búsqueda compleja que vaya incorporando paulatinamente cada uno de los términos para saber en qué momento hay o no posibilidades de recuperar información.

2.5.7 OTROS RECURSOS Y PROCEDIMIENTOS DE EXHAUSTIVIDAD

Como se ha visto, la mayoría de los operadores estudiados tienen como función el incrementar la tasa de precisión de las búsquedas, pero existen también otras técnicas o procedimientos que tienen como finalidad el hacer más exacta la recuperación de información, las detallaré a continuación:

a) Truncamiento

En la mayoría de las bases de datos e inclusive en INTERNET se puede utilizar un asterisco, de forma que en la ecuación se escribe sólo la raíz del término, pero la búsqueda se efectúa sobre todo el conjunto de voces que comiencen con la misma.

ejemplo: ADULT* = ADULT, ADULTS, ADULTHOOD, ADULTERY, ADULTERIES, ETC.

b) Uso del índice

Igualmente, la mayoría de las bases de datos llevan consigo un índice, el cual es de gran ayuda para seleccionar los términos que figuran en los registros de la base de datos, excepto los pertenecientes a campos limitados.

c) Uso de campos específicos de las bases de datos

Para aumentar la eficacia de la recuperación es necesario conocer también en detalle la estructura de cada una de las bases de datos con que se opere.

CAPÍTULO 3.

ESTRATEGIAS DE BÚSQUEDA

Las estrategias de búsqueda son técnicas que se emplean para buscar información de manera efectiva, que conjugan diferentes elementos pertinentes para poder llevar a cabo una recuperación de la información de forma exitosa. En este capítulo se abordarán varios cuestionamientos para poder realizar una estrategia de búsqueda.

3.1 ¿CÓMO DETERMINAR EL TEMA ESPECÍFICO A BUSCAR?

Éste es el paso más importante cuando se inicia una búsqueda de información, a este paso comúnmente se suele dar el nombre de "planeación". La planeación es un tema muy importante en el ámbito administrativo pero, en todos los ámbitos, es pieza clave para cualquier proceso. Con el término "planeación" se denomina el proceso de fijar el curso concreto de acción que ha de seguirse, estableciendo los principios que habrán de orientarlo, la secuencia de operaciones a realizar y las determinaciones de tiempos y números, necesarios para su realización²⁸. Todo control es imposible si no se compara con un plan previo. Sin planes, se trabaja a ciegas. Por todo esto es importante la planeación y más aún en el ámbito informático. El proceso de planear la búsqueda comprende:

²⁸ Reyes Ponce, Agustín. Administración de empresas, teoría y práctica. México: Limusa, 1994. p. 165

3.1.1 DEFINIR UN TEMA ESPECÍFICO PARA LA BÚSQUEDA

Se pretende conseguir una comprensión clara del tema que se está investigando en las bases de datos, este es un aspecto sumamente importante. La persona que está buscando la información debe tomar en cuenta los siguientes aspectos:

- Establecer un contacto personal con el investigador que solicita la búsqueda.
- Asegurarse de entender el tema de la investigación.
- Conseguir que el investigador explique todos los detalles técnicos de la búsqueda.
- Buscar el tema en un libro de referencia general si no está claro aún.
- Conversar con el investigador acerca de la estrategia de búsqueda y obtener su ayuda para seleccionar la terminología.
- Tratar de que el investigador comente sus resultados a medida que vayan apareciendo en la terminal.

Ejemplo: tema general = Inteligencia Artificial

tema principal = Sistemas expertos

tema específico = Diseño de sistemas expertos en el área médica

3.1.2 SELECCIONAR LOS TÉRMINOS DE BÚSQUEDA

Es muy importante elegir cuáles serán las palabras o términos que se utilizarán en la búsqueda de la información, los términos pueden estar presentados en diferentes formas:

a) Palabras clave: se utilizan términos aislados, se agrupan mediante el uso de operadores lógicos y booleanos.

Ejemplo: expert AND systems

b) Frases: cuando se utilizan palabras compuestas (vocabulario controlado), que se pueden obtener mediante el uso de thesaurus.

Ejemplo: expert-systems

Si se utiliza cualquiera de estas dos formas se obtienen resultados satisfactorios. Cabe mencionar que es recomendable utilizar palabras claves aisladas cuando aun no se tiene una estrategia de búsqueda definida y no se conocen los términos de búsqueda que pueden definir exactamente el tema.

Hacer una búsqueda con palabras claves nos lleva a resultados cuyos registros incluyen campos que contienen frases descriptivas, que al ser utilizadas posteriormente llevarán a resultados más específicos. O bien, cuando se conocen los términos, se pueden utilizar thesaurus para definir el vocabulario controlado antes de iniciar una búsqueda.

El éxito de la búsqueda dependerá de dos factores:

1. El cuidado que se tenga en seleccionar y describir las palabras claves o frases descriptoras.
2. De la elección de la base de datos a utilizar.

Estos dos factores son fundamentales para determinar la calidad de la información recuperada.

Las palabras clave son muy importantes para la búsqueda y recuperación de información, por lo que, la selección que se haga ellas deberá encajar con la naturaleza de la base de datos que se utilice; por ejemplo, no es recomendable utilizar palabras muy técnicas en una base de datos no técnica, o preguntar por la información de los 10 últimos años en una base de datos que contiene información reciente solamente.

3.2 PREGUNTAS MÁS IMPORTANTES AL INICIAR UNA INVESTIGACIÓN

Una recomendación muy útil es reflexionar antes de hacer la búsqueda de información y tomar en cuenta ciertos aspectos como los siguientes:

a) ¿Cuál es el tema principal de la búsqueda o investigación?

Para poder definir fácilmente el tema de la búsqueda correctamente es necesario preguntarse lo siguiente:

- ¿Cómo se puede definir el tema de una manera clara en una sencilla oración? o ¿cuál sería el título del artículo perfecto?.

Por ejemplo, si se necesitara información en alguna base de datos norteamericana acerca de dietas bajas en colesterol, el título perfecto en inglés sería: "Low-cholesterol diets".

b) ¿Qué palabras deberá tener un artículo para que se pueda considerar que es relevante con respecto al tema principal?

En este apartado es recomendable hacer un esquema por escrito que con el ejemplo anterior quedaría de la siguiente manera:

Tema principal: **Low cholesterol diets**

Palabras clave:

CONCEPTO 1

cholesterol

AND

CONCEPTO 2

lower

**AND
OR**

CONCEPTO 3

diets
low
reduce
reduction
controls
controlling

c) ¿El tema de la búsqueda amerita utilizar vocabulario científico, técnico o requiere del uso de sinónimos?

En ocasiones, se tiene la base de datos adecuada en la que se sabe que está la información que se requiere, sin embargo, frecuentemente se sabe qué palabras son las adecuadas o simplemente se escriben palabras o frases que no van de acuerdo al vocabulario de la base de datos, ya que se trata a veces de una base de datos científica en la que se está empleando términos de búsqueda coloquiales o al revés. También se requiere de sinónimos y no se emplean y por ello no se encuentra la información luego entonces se piensa que en cierta base de datos no hay información. En conclusión, se recomienda examinar qué tipo de vocabulario es el que se empleará en la base de datos y si es factible el uso de sinónimos.

d) ¿Todos los términos requieren de estar en el registro localizado para considerarlo como relevante?

Actualmente, la mayoría de las bases de datos tienen una pantalla de recursos accesibles y hacen que el proceso de búsqueda sea cada vez más simple, pues desde la primera fase de búsqueda se pueden ingresar los conceptos para que se

busquen en un campo específico; por ejemplo, que busque el término de "VIH" en el campo o registro "SUBJECT", y así, está buscando la información en donde se requiere encontrarla. Pero, si un tema es muy restringido y hay poca información acerca de él es necesario saber todo lo que existe del tema y por ello no se necesita que esté en un determinado campo o registro para considerarlo relevante, esto es algo que habría también considerar desde un principio.

e) ¿La información deberá estar limitada por alguna característica importante como área geográfica, población, industria o actividad económica?

Esto, aunque parece muy sencillo es algo que en muchas ocasiones se pasa por alto y se consultan todas las bases de datos que tratan sobre determinado tema y olvidando que hay que ver específicamente cada base y cuestionar al investigador o a la persona que va a buscar la información, si necesita algo en particular del tema en un determinado país, estado o tal vez quiera información de algún indicador en especial o visto desde alguna u otra perspectiva. Si se tiene en cuenta este punto de vista y se le pregunta al investigador acerca de su tema, es muy probable que al entender su necesidad de información se le puedan brindar los datos idóneos para su investigación.

f) ¿Existe algún término que no desee que aparezca en la búsqueda?

En ocasiones cuando se investiga de un tema se encuentra información que no se necesita porque no tiene nada que ver con lo que se está requiriendo, es necesario delimitar y teclear un NOT o NO para que no se recupere lo que no es necesario;

Ejemplo: si se quiere investigar sobre el tema de SISTEMAS DE INFORMACIÓN y se busca bajo esos términos seguramente nos dará (entre otras cosas) las escuelas que imparten clases de SISTEMAS DE INFORMACIÓN y entonces se tendría que teclear NOT PLANES DE ESTUDIOS.

g) ¿Existe alguna limitación con respecto al idioma en que se encuentra la información?

De antemano hay que ver en qué idioma se encuentra la información para ver si la base de datos es la adecuada según las necesidades del usuario.

h) ¿Es necesario limitar la información a una fecha o un rango de fechas?

Siempre es obligatorio preguntar al investigador si requiere información reciente o histórica.

i) ¿Se requiere hacer una búsqueda intensiva en todas las bases de datos del tema o limitarla a las bases más comunes?

Saber si se necesita información básica, introductoria, extensa o vista desde diversos puntos, esto es necesario conocerlo mediante una reflexión para saber qué es lo que requiere el investigador.

3.3 MÉTODO PARA INSTRUMENTAR UNA ESTRATEGIA DE BÚSQUEDA EFECTIVA.

Aunque las búsquedas de información son todas diferentes, existen ciertos puntos que son iguales en todas las estrategias de búsqueda. A continuación se citan los más básicos para poder llevar a cabo una búsqueda:

3.3.1 PROCESO PARA DESARROLLAR UNA BÚSQUEDA

a) Examinar y planear las necesidades de información

En esta primera etapa, es importante definir el tópico o tema exacto para lo cual se necesita la información, para ello se requiere:

- a) Escribir la búsqueda en términos claros.
- b) Pensar cuáles serán los términos que se van a usar para formular adecuada y claramente la búsqueda. Se tiene que considerar qué diferentes conceptos o ideas establecen el tópico completamente. La mayoría de los tópicos de búsqueda consisten en dos o más conceptos, es recomendable auxiliarse con diccionarios generales, de términos especializados, de sinónimos y de antónimos.

Ejemplo: Formular el tema de búsqueda de la siguiente manera:

“Encontrar artículos acerca del uso de INTERNET como herramienta en la investigación superior”

b) Elegir los términos de búsqueda

Este planteamiento ayuda a establecer claramente cuál es el objetivo de la búsqueda, para qué o con qué términos se describe la búsqueda.

Existen dos diferentes conceptos a ser buscados: INTERNET e investigación superior. La búsqueda puede ser tanto por “palabras individuales” (key words) o por frases o temas (subjects).

c) Crear una estrategia de búsqueda

Para crear una estrategia de búsqueda se pueden seguir los siguientes pasos:

1. Determinar cuáles son los conceptos que están presentes dentro del tema o tópico de la búsqueda.
2. Definir cuál es el concepto más relevante y buscarlo primero.
3. Modificar la búsqueda utilizando los operadores lógicos que ayudarán a tener resultados más certeros en la estrategia de búsqueda.
4. No elegir términos generales como “planificación” o “educación”.
5. Si se utiliza algún índice se tienen que elegir frases específicas que contengan esos términos, los cuales deben ser pertinentes al tema como podrían ser: “educación sexual” o “planificación familiar”.
6. Si se introducen directamente palabras o frases que tienen raíces comunes como “programa” o “enseñanza”, evitar el uso de truncadores.
7. Procurar no usar sufijos o palabras que confundan o desvíen el significado de los términos o palabras.
8. Muchas veces un concepto o término puede ser descrito con varios sinónimos o términos alternativos que son similares en el significado. Es importante que al seleccionar los términos de búsqueda e iniciar el proceso de la misma se tengan en cuenta otros términos que se puedan emplear para obtener más y mejores

resultados, por lo que el uso de los sinónimos sirve para aumentar la cantidad y la calidad de los resultados de la búsqueda.

9. Si se recuperan muchos registros, se tiene que limitar la extensión de la búsqueda eliminando los términos irrelevantes, hasta que los resultados de la búsqueda sean aceptables.
10. En caso de que se recuperen cero resultados, se debe tratar de usar términos de búsqueda más generales y no tan específicos.

En términos generales estos puntos son recomendaciones para formular las estrategias y tener éxito en éstas. El proceso de la búsqueda de información es un proceso de comprensión y análisis, que tiene que hacerse con detenimiento, siempre pensando que se puede conseguir la información que se necesita si se emplean correctamente los recursos disponibles.

3.4 SELECCIÓN DE LAS BASES DE DATOS

Después de tomar en cuenta todas las recomendaciones anteriores en el proceso para desarrollar una búsqueda y habiendo elegido, examinado y planeado las necesidades de información, elegido los términos de búsqueda y creado una estrategia de búsqueda, entonces ahora hay que asegurarnos y hacer hincapié en la elección adecuada de la o las bases de datos adecuadas que nos permitan recuperar la información que deseemos de manera exitosa.

3.4.1 VARIACIONES DE LAS BASES DE DATOS

Existe una gran diversidad en bases de datos, las cuales pueden ser clasificadas de la siguiente forma:

- **Por el contenido:** negocios, legal, medicina, tecnología, etc.
- **Por el tipo:** citas bibliográficas, resúmenes, textos completos, directorios.
- **Por la presentación:** textos, imágenes, tablas estadísticas.
- **Por el nivel técnico de los datos que contiene:** reportes de investigación, artículos de revistas o periódicos.
- **Por el tipo de datos que contiene:** actualizados, publicados, no publicados.

3.4.2 POR QUÉ ELEGIR ENTRE UNA BASE DE DATOS Y NO OTRA

La elección de una base de datos determinada depende del tipo de información que el usuario requiera. Muchas veces los usuarios tienen una idea genérica o “vaga” de lo que quieren localizar; por ejemplo, si un usuario requiere información acerca de un determinado país se puede encontrar con información: bibliográfica, numérica, política, financiera, noticiosa, algún directorio del país, etc. Y de acuerdo a eso es que se elegirá entre las bases de datos propuestas la que se adapte más a sus necesidades de información, por lo tanto, si quiere información estadística no pedirá una base de datos que contenga textos completos, pues sabe que no encontrará en ella la información idónea para cubrir sus necesidades de información.

Es necesario verificar si el contenido de las bases de datos es de un tipo específico: si es de tipo histórico, si se actualiza cada determinado tiempo, temas principales que abarca, sus proveedores, etc. Además de tener en cuenta la cobertura geográfica de la información, por ejemplo si es nacional o internacional.

3.4.3 SUGERENCIAS PARA SELECCIONAR UNA BASE DE DATOS

En ocasiones se necesita pedir o directamente consultar una base de datos, por lo cual es indispensable que se tomen en cuenta algunas cuestiones generales para saber qué base de datos consultar y así poder hacer el proceso de búsqueda y recuperación de información más rápida, fácil y exitosa.

A continuación se dan algunas sugerencias para la selección de las bases de datos:

1. Realizar la búsqueda de información primero en las bases de datos mayores y más apropiadas.
2. Las bases de datos menores y más especializadas ofrecen ventajas cuando se necesitan unas cuantas referencias muy pertinentes más que una investigación exhaustiva.
3. El valor de una base de datos depende en gran parte de la literatura que contiene. Se recomienda analizar las fuentes de la misma.
4. Es importante contar con todas las herramientas auxiliares de búsqueda posibles. El uso de thesaurus nos permite llevar a cabo la búsqueda de forma efectiva ya que la eficacia del proceso de la recuperación de información depende de la igualdad de los términos usados para describir una necesidad de información con los términos usados en la representación de documentos. El éxito de los usuarios o investigadores que acuden a realizar una búsqueda puede ser mejorado grandemente aumentando el número de los términos de la búsqueda por concepto,

esto especialmente cuando es ayudado por las herramientas como los thesaurus o algún otro control o listas del vocabulario.²⁹

²⁹ Harrison, Lauren Deborah. *The impact of thesauri on information retrieval (query formulation)*. Rutgers: New Brunswick, 1998. p. 52.

CAPÍTULO 4

EL INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

4.1 EL SISTEMA ITESM

El Instituto Tecnológico y de Estudios Superiores de Monterrey fue creado en Monterrey en 1943. El grupo de empresarios que fundó el Tecnológico de Monterrey se constituyó en una asociación civil denominada "Enseñanza e Investigación Superior, A.C." (EISAC). Posteriormente se fundaron asociaciones similares para iniciar la operación de los campus establecidos fuera de la ciudad de Monterrey.³⁰

A través de estas asociaciones civiles el Instituto cuenta con más de 426 consejeros, que son distinguidos empresarios mexicanos comprometidos con el avance y calidad de la educación superior en México.

Actualmente el proyecto original se ha transformado en un sistema nacional, establecido en 25 ciudades del país, en el que diariamente 80 mil personas se dedican a la vida académica con elevados índices de calidad. Cuenta también con una Universidad Virtual, mediante la cual el Instituto tiene una presencia importante en Latinoamérica.

³⁰ Mendirichaga, Rodrigo. El Tecnológico de Monterrey: sucesos, anécdotas, personajes. México: Castillo, 1982. p.54.

El Instituto ofrece 31 carreras profesionales, 37 maestrías y 7 doctorados. Los programas académicos del Tecnológico de Monterrey están acreditados en los Estados Unidos de Norteamérica por la Asociación de Universidades y Escuelas del Sur (Southern Association of Colleges and Schools) y se renuevan periódicamente para mantener una relación estrecha con el mercado de trabajo, así como para integrar la nueva tecnología a las diferentes áreas académicas centrando su atención en el espíritu emprendedor, en los valores y la ética profesional.

4.1.1 MISIÓN DEL ITESM

La definición proyectiva de la Misión del Tecnológico de Monterrey hacia el año 2005, fue resultado de un amplio proceso de consulta entre diversos grupos académicos a nivel nacional.

En esta consulta se identificaron los retos más importantes que tiene el país y se definieron cuáles representaban oportunidades para que el Tecnológico enfocara tanto la labor formativa que realiza con los alumnos, como los de investigación.

Los retos en los que basa su Misión son:

- Creación de empleos,
- Apoyo a la competitividad,
- Definición de estrategias para el desarrollo nacional, regional y sectorial, la participación en el proceso de democratización y el mejoramiento de la educación

Teniendo en cuenta estos retos y el perfil de los alumnos, profesores, egresados, consejeros y la operación misma, así como la naturaleza del proceso enseñanza-aprendizaje, la investigación y la extensión la institución tienen como

Misión:

- Formar personas comprometidas con el desarrollo de su comunidad, para mejorarla en lo social, en lo económico y en lo político.
- Hacer investigación y extensión relevantes para el desarrollo sostenible del país.³¹

El Instituto tiene su propio programa de becas y préstamos educativos que beneficia al 33% de alumnos de preparatoria, profesional y posgrado. El alumnado está concentrado mayormente en las áreas de ingeniería, computación y administración.

La población estudiantil al nivel de posgrado es de más de 6,300 alumnos. El Instituto ofrece, además, preparatoria bilingüe y bicultural así como otros programas educativos.

El Sistema del Tecnológico de Monterrey está interconectado mediante redes computacionales y promueve masivamente entre sus profesores y alumnos el uso de la telecomunicación en la actividad académica.

El Tecnológico de Monterrey cuenta en todo el Sistema con 15,500 computadoras, de las cuales 8,800 están a disposición de los alumnos, lo que da una proporción de 8

³¹ Instituto Tecnológico y de Estudios Superiores de Monterrey. Principios, misión, organización y estatuto general del sistema Instituto Tecnológico y de Estudios Superiores de Monterrey. México: ITESM, 1986. p. 43.

alumnos por computadora. Son cada vez más numerosos los cursos que en el proceso de enseñanza-aprendizaje incluyen el uso de tecnologías computacionales tales como "home pages" y "group ware".

El Instituto entrelaza sus campus por medio de una red de comunicaciones con 5 accesos de alta velocidad a Internet y una red satelital.

La Universidad Virtual del Sistema del Tecnológico de Monterrey imparte clases por satélite a instituciones educativas de México, Centro y Sudamérica.³²

4.2 EL CAMPUS CIUDAD DE MÉXICO DEL ITESM

Las actividades académicas del Instituto Tecnológico y de Estudios Superiores de Monterrey en la ciudad de México comenzaron en septiembre de 1973 en el antiguo edificio de la Escuela de Graduados en Administración, ubicado en la avenida Fray Servando Teresa de Mier. Conscientes de la creciente demanda de educación media y superior que existía en el sur de la ciudad, a partir de agosto de 1990, el Campus Ciudad de México comienza sus actividades en las instalaciones del área de Tlalpan.

Actualmente, el Campus Ciudad de México ofrece un Programa Doctoral en Administración, ocho Programas de Maestría, más de quince programas de educación continua para ejecutivos, dieciséis carreras profesionales y dos programas de preparatoria. Atiende a más de 10,000 estudiantes. El 28 por ciento se encuentra en

32 López Lopez, Juan Eduardo. Camino a la excelencia: la transformación del cuerpo docente del Sistema Tecnológico de Monterrey de 1989 a 1995. México: ITESM, 1995. p. 12.

preparatoria, el 64 por ciento en profesional y el resto, compuesto por más de 700 alumnos, pertenece a los programas de posgrado.

Continuamente se ofrecen a los profesores cursos de actualización y especialización dentro del área en la que imparten sus clases, adicionalmente a los cursos de desarrollo de habilidades docentes.

4.3 LA BIBLIOTECA DEL ITESM CCM

La biblioteca del Instituto Tecnológico y de Estudios Superiores de Monterrey, campus Ciudad de México es el corazón de la vida académica y de investigación del Campus.

Fue inaugurada oficialmente en octubre de 1993, ya que anteriormente se encontraba en un salón del segundo piso de un edificio llamado Aulas I, que es en donde se imparten las clases a los alumnos de nivel profesional; posteriormente se ubicó en el primer piso del edificio de Oficinas I, en el cual se llevan a cabo los procesos administrativos del Campus; actualmente cuenta con cinco niveles y es por ello, uno de los edificios más importantes del Campus por que ha contribuido y contribuye a formar individuos competitivos y está a la vanguardia internacional.

4.3.1 DATOS GENERALES DE LA BIBLIOTECA DEL ITESM CCM

ÁREA DE CONSTRUCCIÓN DONDE SE ENCUENTRA LA BIBLIOTECA	8,445 METROS CUADRADOS
Área que corresponde a la Biblioteca	6,046 metros cuadrados
Lugares de estudio; en donde dos usuarios estudian de manera individual cómodamente	1,550
Computadoras para consulta e investigación	80
Bases de datos en Red, CD-ROM, y Disco duro.	224
Promedio diario de usuarios en la Biblioteca	10,000
Número de Instituciones con las que mantiene préstamos interbibliotecarios	55
Volúmenes totales de Libros	221,366
Títulos de revistas y periódicos	519

La biblioteca del campus así como la división de Informática, de la que forma parte, tiene como misión proporcionar servicios y tecnologías de información a la vanguardia para contribuir a la calidad en los procesos de enseñanza-aprendizaje, investigación y de administración que se desarrollan en el Campus Ciudad de México³³ y además con esto tiene el compromiso de fomentar entre los usuarios el desarrollo de los valores y habilidades planteados en la Misión del Sistema ITESM hacia el 2005³⁴, como son:

33 Instituto Tecnológico Y de Estudios Superiores de Monterrey, Campus Ciudad de México. *Guía de Informática*. México: ITESM, 2000. p. 2

34 Misión del Sistema ITESM hacia el 2005: Formar personas comprometidas con el desarrollo de su comunidad, para mejorarla en lo social, en lo económico y en lo político, y que sean competitivas internacionalmente en su área de conocimiento. La Misión incluye hacer investigación y extensión relevantes para el desarrollo sostenible del país.

- honestidad,
- responsabilidad,
- respeto a la dignidad de las personas,
- aprendizaje por cuenta propia,
- capacidad para identificar y resolver problemas,
- trabajo en equipo,
- uso eficiente de la informática y telecomunicaciones
- sentido de arraigo y pertenencia al instituto.³⁵

4.3.2 MISIÓN DE LA BIBLIOTECA

La misión que corresponde a la biblioteca del ITESM CCM es:

Satisfacer las necesidades de información de la comunidad del Campus Ciudad de México a través de la promoción de la lectura, la actualización constante de las fuentes de información con la incorporación de tecnologías de vanguardia para la formación de personas internacionalmente competitivas en su especialidad.³⁶

4.3.3 OBJETIVOS DE LA BIBLIOTECA

- Detectar oportunamente las necesidades de los usuarios y desarrollar las colecciones de acuerdo al estado del arte de cada disciplina.
- Garantizar la calidad y el fácil acceso de la comunidad del Campus Ciudad de México a los materiales impresos, audiovisuales y en formatos no convencionales

³⁵ Instituto Tecnológico Y de Estudios Superiores de Monterrey, Campus Ciudad de México. Manual de biblioteca. México: ITESM, 1998. p. 1.

³⁶ Ibidem p. 3

(bases de datos en línea, en disco, etcétera) por medio de la organización adecuada y consistente de las colecciones.

- Orientar a la comunidad sobre el uso y mejor aprovechamiento de los recursos disponibles en la Biblioteca.
- Contribuir con su material al desarrollo de los programas académicos, de investigación y de extensión que incidan en la obtención y difusión del conocimiento.
- Estimular el interés por la lectura y el conocimiento
- Promover el uso efectivo y constante de la Biblioteca en las actividades relacionadas con la docencia y la investigación.³⁷

4.3.4 POLÍTICAS GENERALES DE BIBLIOTECA*:

Los documentos oficiales mencionan como políticas los siguientes puntos:

- Presentar la credencial vigente que lo acredite como miembro de la comunidad del ITESM Ciudad de México para solicitar cualquier servicio.
- Esta prohibido fumar e introducir alimentos en cualquier área de la Biblioteca.
- El usuario que se sorprenda subrayando, extrayendo información sin autorización o haciendo mal uso del material será acreedor a la sanción que determine un comité especialmente integrado para analizar el caso y su decisión será irrevocable.
- La persona que se encuentre mutilando el material de la Biblioteca será expulsado del Sistema ITESM.

³⁷ Ibidem p. 7

*A pesar de que muchos de ellos corresponden más a la formulación de un reglamento que a políticas.

- Se prohíbe deteriorar o hacer mal uso de las instalaciones (rayar las mesas, cambiar el mobiliario de lugar, dañarlo, modificarlo, entre otros.)
- Cualquier punto que amerite sanción y no se mencione aquí, será analizado por la dirección de la Biblioteca y las autoridades correspondientes.³⁸

Lo anterior a pesar de que parecería parte de un reglamento, como lo menciono en los manuales e incluso en la página de internet de la biblioteca aparecen como políticas.

4.4 EL CENTRO ELECTRONICO DE INFORMACIÓN “CEI”

El Centro Electrónico de Información es un servicio de consulta a bases de datos que concentra información tanto referencial como de texto completo, programas multimedia y libros electrónicos en diversos formatos. El acceso puede ser a bases de datos remotas o locales.

El CEI es una sección moderna y actualizada de la Biblioteca en la cual se puede encontrar información de todas las áreas del conocimiento tanto en discos compactos como en bases de datos y sistemas en línea.

4.4.1 MISIÓN DEL CEI

Aunque no está publicada la misión del CEI, ni en documentos que esten a punto de publicarse se plantea como misión la siguiente:

- Satisfacer las necesidades de información del campus, tanto de profesores como de alumnos y a todo el personal del ITESM CCM con un servicio amable, de calidad y profesional mediante la actualización continua de los sistemas disponibles.

- Desarrollar las habilidades en el manejo de bases de datos que no solamente permitan usar el CEI sino también la búsqueda en cualquier otra fuente.

4.4.2 OBJETIVOS DEL CEI

- Ofrecer información actualizada y completa para todas las áreas académicas del campus (desde preparatoria hasta doctorado), con acceso rápido y fácil para el usuario.
- Apoyar a los usuarios en la búsqueda y obtención de información.
- Mantener un vínculo con la comunidad del ITESM CCM para que así se convierta en un centro facilitador de información.

4.4.3 RECURSOS DEL CEI

- Se cuenta con 43 PC's de alta capacidad.
- Un servidor HP NET SERVER*
- Dos JUKE BOX* con 14 unidades cada uno donde se exponen 28 CD's remotamente para consulta de cualquier computadora conectada en el Campus Ciudad de México.
- Acceso a los servicios de INTERNET para la búsqueda de información pública y privada.
- 240 títulos de bases de datos, como CD-ROM, de acceso vía INTERNET o instaladas en disco duro.

38 Instituto Tecnológico Y de Estudios Superiores de Monterrey, Campus Ciudad de México. Manual de biblioteca. México: ITESM, 1998. p. 8.

* HP NET SERVER que tiene capacidades de memoria suficientes para manejar múltiples accesos, aproximadamente 300 usuarios, de forma rápida y simultánea para acceder a la información.

* JUKE BOX es un dispositivo que sirve para controlar y administrar múltiples unidades de discos compactos mediante unidades lectoras conectadas a un servidor.

- 3 Impresoras.
- Un Escáner de imágenes.

4.4.4 SERVICIOS DEL CEI

Los servicios del CEI se clasifican en tres categorías:

- Bases de datos en línea
- Bases de datos en discos compactos
- Bases de datos en disco duro

4.4.5 USUARIOS DEL CEI

Los alumnos, maestros e investigadores y personal del ITESM CCM se consideran como usuarios internos y tienen acceso a este centro sin ninguna restricción, así como investigadores externos cuyas instituciones de origen tengan convenio con el Centro Electrónico de Información.

4.4.6 VENTAJAS DEL CEI

Las ventajas que ofrece el CEI a los usuarios para sus investigaciones o consultas con respecto a otras áreas de la biblioteca pueden resumirse en las siguientes:

- El usuario llega con una pregunta y sale con una respuesta en un mínimo de tiempo.
- El usuario tiene acceso a la información actualizada.
- El usuario reduce sus costos de investigación.

- El CEI es el único lugar del Campus donde se puede encontrar información estadística y en general actualizada de cualquier tema en formato electrónico.
- Siempre hay computadoras disponibles para las búsquedas de información.
- Es posible acceder a las bases de datos de las dos Bibliotecas Digitales existentes (la del Campus y la del Sistema).
- La información es amplia y variada.
- También se encuentra un centro de digitalización para profesores en el cual pueden dejar sus apuntes, presentaciones, resúmenes y cualquier material de apoyo para sus clases, y así, se digitalizan sus apuntes y los alumnos a través de la red del campus pueden tener acceso y consultar o bien, grabar la información si así lo requieren.

El CEI cuenta con 43 computadoras que permiten el acceso a la información sobre:

Comercio y Administración,

Legislación y Normatividad,

Economía y Finanzas,

Geografía Económica y Censos,

Electrónica y Computación,

Normas Industriales,

Educación,

Directorios,

Información que aporta el INEGI,

Revistas de política e interés general,

Diarios,

Multimedia: cine, astronomía, música, historia, anatomía, entre otros.

4.4.7 DESCRIPCIÓN DE LAS PRINCIPALES BASES DE DATOS EN LÍNEA DEL CEI

El CEI ofrece el servicio de conexión remota a bancos de información, los cuales pueden ser consultados a través de computadoras conectadas a las redes nacionales e internacionales. Estas contienen todo tipo de información publicada en cualquier parte del mundo. Por ejemplo:

- **Infosel Financiero:** Es una base de datos en la que se tiene el monitoreo de las actividades bursátiles de la Bolsa Mexicana de Valores a tiempo real y donde además se pueden consultar noticias, series de tiempo de las empresas mexicanas que cotizan en la Bolsa de Valores de México.
- **Infosel Legal:** Contiene información de todas las leyes, reglamentos, decretos y acuerdos del Distrito Federal y algunos estados de la República Mexicana, además también abarca la jurisprudencia y resoluciones de las Salas.
- **IF PLUS:** Contiene información de los artículos de los periódicos Reforma y El Norte de 1994 hasta el día en curso.
- **Infolatina:** A través de este servicio, se tiene acceso al acervo histórico de más de 150 fuentes de información de México para consultar simultáneamente en todas ellas los temas de interés. Contiene más de 800,000 documentos con información financiera, jurídica, política, legal o de

negocios que proviene de fuentes mexicanas. Además, permite obtener textos completos de documentos como periódicos, revistas, estudios, estadísticas, publicaciones del gobierno, información de la Bolsa Mexicana de Valores y directorios especializados. El periodo que abarca la información depende de la fuente de la que se trate desde 1980 a 1999 ya que diariamente se actualizan las fuentes de información.

- **STAT-USA:** Comprende información sobre demografía, economía, noticias sobre el tema económico, comercio exterior, información financiera de empresas transnacionales, el perfil de los países, estudios demográficos y de mercado, información especializada por producto y país de origen y oportunidades de negocios a nivel internacional, oferta y demanda de productos, entre otros temas.
- **The World Factbook:** Base de datos en donde se pueden consultar diversos temas de interés de todos los países del mundo, tales como: cultura, población, industria, política, economía, información gubernamental; además permite consultar directamente los periódicos y revistas propias del país, así como tener acceso a las direcciones electrónicas de las diferentes instituciones y asociaciones nacionales.
- **Lanic:** Esta base se puede comparar con la anterior sólo que la información se limita a los países de Latinoamérica.
- **Enciclopedia Britannica Online:** Por ser un servicio en línea, la enciclopedia está en constante actualización. Diariamente se agregan a ella nuevos artículos, revisiones y páginas de Internet seleccionadas. Se puede explorar información desde once bases de datos. También se obtiene acceso

a referencias de libros, accesos por categoría a la información de la enciclopedia (spectrum), artículos clásicos de ediciones anteriores de la enciclopedia y una serie de artículos especiales.

- **ProQuest Applied Science & Technology Plus:** Brinda a estudiantes de ingeniería la fuente esencial para un amplio rango de disciplinas. Esta comprehensiva base de datos combina de H. W. Wilson's el índice de Applied Science & Technology con una cobertura completa de las publicaciones líderes en ciencia y tecnología. Incluye más de 450 títulos que presenta fichas bibliográficas y abstracts de 50 a 100 palabras para cada artículo. Además, el usuario puede ir de la bibliografía a la imagen completa para más de 100 de las mejores, que incluyen gráficos, diagramas e ilustraciones que son tan vitales en la literatura científica y tecnológica. Dentro de los temas que cubre esta base de datos se pueden encontrar: aeronáutica, ingeniería química, tecnología en computación y sus aplicaciones, ingeniería eléctrica, alimentación e industria, ingeniería mecánica, ingeniería nuclear, física, telecomunicaciones, transportación y más. ProQuest Applied Science & Technology Plus se actualiza cada semana e incluye bibliografías desde 1994 hasta la fecha y texto completo con imágenes de 1997 a la fecha.
- **ProQuest Telecommunications:** Contiene los principales diarios especializados en Telecomunicaciones en texto completo, da a los investigadores fácil acceso en línea a los artículos en texto completo que ellos necesitan, fácilmente entregables y con derechos pagados, por medio de ProQuest Direct. ProQuest Telecommunications se constituye de la mejor

selección de diarios que existen en línea. Algunos de ellos son: America's Network, Telecommunications, Telephony, European Broadband Networking News, Journal of Broadcasting & Electronic Media. Se actualiza diariamente. Las citas, resúmenes, y el texto completo ASCII se proveen para cada registro.

- **Digital Dissertations 1997:** Contiene más de 1.5 millones de tesis doctorales. Todas las tesis incluyen el nombre del autor, la fecha de publicación, las fuentes bibliográficas y la facilidad de recuperar cuantas copias sean requeridas ya sea en formato de microfilm, microficha y ahora en formato de Acrobat.
- **Social Science Plus Text:** Establece un nexo entre resúmenes e índices con cobertura de más de 400 títulos de Wilson's Social Science Index a texto completo e imagen completa de más de 200 de esos títulos. Las disciplinas cubiertas incluyen estudios de adicción, gerontología, economía, relaciones internacionales, estudios urbanos y otros. La cobertura del índice comienza en 1994. La base de datos se encuentra disponible en CD-ROM, con cobertura desde 1989.
- **EIU International Business Newsletters:** Contiene información sobre los últimos desarrollos de negocios en más de 180 mercados emergentes a nivel mundial. Es la colección completa desde 1993 a la fecha de la publicación.
- **ProQuest Medical Library:** Combina MEDLINE indexing con imágenes completas para más de 100 títulos médicos claves. ProQuest Medical Library lleva a los usuarios desde la cita bibliográfica a la imagen completa en una sencilla operación. La imagen completa es esencial en campos clínicos,

donde mapas, diagramas, fotografías, y otros materiales visuales son una parte integral del texto. ProQuest Medical Library da formato completo, alta resolución de imágenes en blanco y negro en la versión en CD y otras muchas a color en la versión conectada. En esta base es fácil de encontrar la información que los usuarios quieren, con un exacto indexado para cada título de imagen completa. Los diarios cubren todas las especialidades importantes de salud pública, incluyendo enfermería, pediatría, neurología, farmacología, cardiología, fisioterapia, y otros. Entre estos diarios están: The New England Journal of Medicine, JAMA: Journal of the American Medical Association, Lancet, Archives of Surgery, British Medical Journal, Nursing, Pediatrics, y otros más.

- **Books in Print:** Provee acceso a alrededor de 1.5 millones de registros y 50,000 contactos con editores, publicistas y distribuidores. Ofrece acceso inmediato a información de la industria editorial del libro en Estados Unidos e incluye también acceso a los libros fuera de impresión.
- **General Science PlusText:** Provee acceso confiable y rápido a la información científica con los patrones que los investigadores necesitan en una gama amplia de disciplinas, incluyendo la astronomía, biología, ciencias de la tierra, genética, medicina, alimentación, y más. La base de datos combina índices de 169 títulos de Wilson General Science Abstracts Index, con texto completo y/o imagen completa para más de 50 títulos claves. La cobertura del índice comienza en 1994.
- **General Periodicals:** Da la cobertura en un rango de revistas fácilmente accesibles; revistas populares, publicaciones clave en negocios con la

imagen completa de cientos de las más importantes y populares revistas. Presenta en texto completo desde 1988 y tiene un retrospectivo total a 1986. General Periodicals Research contiene índices y abstracts de más de 1,800 publicaciones más 6 de los meses más recientes de El New York Times y The Wall Street Journal e imagen completa para más de 450 publicaciones.

- **ProQuest Computing:** Es la selección más grande de diarios que se encontrará en línea, gracias a sus actualizaciones diarias. Entre los 200 principales diarios están: Journal of the Association for Computing Machinery, Communications of the ACM, Byte, Computer Reseller News Computerworld, InfoWorld. La cobertura incluye no sólo revisiones de producto, sino también otro tipo de artículos: Administración y diseño de base de datos; Estrategias de desarrollo de Software; Comercio en Web; LANS y WANS; Intranets e Internet; Administración de Tecnología. Los registros incluyen resúmenes y citas bibliográficas completas. El texto completo que está en código ASCII se provee para el 95% de los títulos.
- **El Business Periodicals:** Provee de la más calificada cobertura en publicaciones disponibles en administración y negocios. Esta base de datos brinda la imagen completa de los artículos más importantes y más populares de la fuente de ABI/INFORM. Contiene el índice y los resúmenes de los artículos de las mejores revistas, además las imágenes completas de por lo menos la mitad de las publicaciones. Con una cobertura retrospectiva a 1971. La imagen completa empieza en 1988, contiene el índice y los resúmenes de más de 1000 publicaciones, incluyendo más de 400 títulos de

revistas en inglés de todo el mundo. La imagen completa está disponible para más de la mitad de los títulos.

- **Internet Securities:** Base de datos que contiene información de diferentes países del mundo en la que los temas que destacan son: noticias de la compañía, los datos financieros de más de 25 mercados que emergen en Asia, América y Europa Oriental. La información va desde una noticia actual y difícil de conseguir hasta información estratégica de las compañías más importantes, industrias, información financiera y de mercado, pronósticos, situación política del país, datos para hacer un análisis macroeconómico detallado. Ofrece una solución conveniente para hacer un negocio con inteligencia. La cobertura comprende América Latina, América Central y Europa Oriental.
- **Industry Net:** Contiene el mercado electrónico en Internet para la industria. De manera rápida, fácil y productiva se obtiene la información importante del producto industrial y de compañías. Industry Net ensambla a una comunidad de más de 550.000 profesionales de la industria y se encuentra actualizada hasta la fecha. Ofrece dos tipos de servicio: de calidad de miembro libre y calidad de miembro superior.
- **Economática:** Base de datos que reúne información de las Bolsas de valores de todos los países de América Latina y Estados Unidos. Además contiene información de las empresas cotizadoras en las principales Bolsas del continente Americano con pronósticos, balances generales, acciones, estados financieros y noticias de la compañía.

4.4.8 DESCRIPCIÓN DE LAS PRINCIPALES BASES DE DATOS EN DISCO

COMPACTO DEL CEI

El CEI cuenta con 180 títulos de discos compactos que van desde enciclopedias y diccionarios hasta bases de datos especializadas como por ejemplo, las bases de la IEEE y de World Wide Standars Service Plus, las cuales son, unas de las más importantes del área de ingeniería. A continuación se dará una breve descripción de las base de datos más importantes o más utilizadas:

- **IEEE/IEE:** Provee información para poder investigar es una sola fuente información del ámbito eléctrico, la literatura de computación e ingeniería. Se puede tener acceso a todas las publicaciones del Instituto de Ingeniería Eléctrica y Electrónica de Estados Unidos (IEEE) y de la Institución de Ingenieros Eléctricos (IEE). Contiene mas de 12,000 publicaciones con mas de 2 millones de imágenes de páginas de 12 años hacia atrás. Rápidamente se encuentra información de fondo, reta hipótesis, valida asumpciones, y genera ideas mientras se va desarrollando nuevos conocimientos acerca de la tecnología. Actualmente cuenta con más de 500,000 artículos de 12,000 publicaciones, incluyendo diarios, conferencias, procedimientos, normas en más de dos millones de imágenes de las páginas completas incluyendo todos los mapas originales, diagramas, fotografías e ilustraciones; el material. Abarca publicaciones desde 1988 hasta el momento actual.
- **Political Risk:** Esta base de datos muestra información actual sobre la mayoría de los países del mundo. Contiene estudios políticos y los pronósticos económicos, analiza corrientes y futuras condiciones para el negocio internacional incluyendo datos actuales, pronósticos a corto y largo

plazo. Además se incluyen algunas gráficas y mapas de los países. Toda esta información es anualmente actualizada y es una eficiente herramienta para diversas áreas como comercio internacional, administración, economía, ciencias sociales y derecho.

- **Moody's Global:** Disco compacto que contiene la información de las 17,700 empresas más importantes del mundo; incluye las empresas mexicanas que cotizan en la Bolsa Mexicana de Valores. La información se encuentra dividida por folders y así se tienen datos de las compañías tales como su historia, dirección, actuales directivos, área de sus negocios o giro empresarial, filiales o subsidiarias que hay en diversos países, balances generales, reportes financieros anuales, noticias trascendentales de la compañía.
- **World Consumer Market:** Disco compacto que reúne información de aproximadamente 256 productos de consumo de 56 países del mundo en donde viene el consumo de estos productos por año y por país o región, los datos abarcan desde 1991 hasta 1996.
- **World Marketing Data:** Disco compacto que abarca 1055 datos estadísticos de negocios y mercados en 209 países.
- **World Wide Standards Service Plus:** Esta base de datos, que comprende varios discos compactos, contiene información necesaria en las ramas de la ingeniería, tales como la petroquímica, ingeniería automotriz, civil, eléctrica, electrónica, instrumentación e industria del transporte y otros temas relacionados en todo el mundo. A través de ella se tiene acceso a más del 90% de los estándares más usados y aceptados en las industrias

internacionales. Los estándares que contiene son de más de 440 Organizaciones de Estados Unidos de Norteamérica.

- **TeleComm Standars Collection:** Contiene acceso a más de 4500 estándares técnicos, de varias sociedades que regulan y controlan la industria de las telecomunicaciones. Los estándares cubren señales de transmisión, construcción de redes de trabajo, protocolos, etc. La colección de TeleComm está integrada por diferentes módulos: símbolos y definiciones, pruebas y mediciones, microondas, fibras, equipos, etc.
- **Indicators of Tariff & Non-Tariff:** Base de datos de la OCDE para hacer un estudio detallado país por país de las barreras arancelarias de todos los países miembros de la OCDE para apoyar exportaciones e importaciones.
- **Bancomext Trade Directory:** Disco compacto que reúne toda la información de todas las empresas mexicanas dedicadas a exportar a cualquier país del mundo. Los datos que contiene son la dirección de la compañía, actividad a que se dedica, los productos que está exportando y hacia dónde, además, dedica una pequeña parte para saber si la empresa también está importando y conocer la fracción arancelaria de cada producto.

A continuación se enumerarán, dentro de esta misma categoría de bases de datos en disco compacto, otras bases que son más conocidas en el área.

- **How realy start your own business:** Disco compacto en el que se describen técnicas de fácil comprensión para iniciar un negocio propio.
- **Computer select:** Información referente a hardware, software y tendencias en informática.

- **El Financiero:** El diario capitalino con todas sus secciones y notas. Incluye tablas desde 1985 hasta 1998.
- **Áreas Metropolitanas (1989-1994):** Información estadística y cartográfica de las principales ciudades del país.
- **Códice 90:** Resultados definitivos del XI Censo General de Población y Vivienda, 1990.
- **Conteo de Población y Vivienda 1995; México (Resumen):** Actualización del Censo de población de 1990 de todos los estados de la República Mexicana.
- **Conteo de Población y Vivienda 1995:** Es la actualización del Censo de población de 1990 de todos los estados de la República, estado por estado.
- **Empleo Urbano:** Incluye datos estadísticos sobre los indicadores y variantes del empleo urbano en México.
- **Natalidad:** Estadísticas sobre la tasa de natalidad de México por zonas y características de la población.
- **Nupcialidad:** Estadísticas sobre matrimonios en México detallado por edades, zonas, etc.
- **Mortalidad:** Estadísticas sobre la tasa de mortandad en México con sus principales causas.
- **Estructura y Dinámica Poblacional:** Estadísticas sobre crecimiento, migraciones y emigraciones en México.
- **Ingreso-Gasto de los Hogares:** Tabulaciones y registros de las estadísticas sobre la relación ingreso-gasto en los hogares mexicanos.

- **Sector Agropecuario (1991):** Resultados definitivos del Censo Agrícola, Ganadero y Ejidal de 1991.
- **Estadísticas Históricas de México:** Compendio estadístico de los indicadores más representativos de México desde 1857 hasta 1970.
- **La política económica de México 1946-94:** Como su nombre lo indica abarca estadísticas económicas de México 1946-1994.
- **Consulta de Información Económica Nacional (CIEN):** 251 indicadores censales de los resultados definitivos del Censo de 1994.
- **La Economía Mexicana en Cifras:** Cifras económicas de México proporcionadas por Nacional Financiera de 1980 a 1994.
- **Cuentas Nacionales de México 1994:** Series anuales y trimestrales de los indicadores económicos de coyuntura.
- **Sistema Autorizado de Información Censal:** Incluye información sobre estructura económica del país, entidades federativas y sus municipios.
- **Censos Económicos 1994. Resultados Oportunos:** Resultados del censo de servicios, económico y manufacturero de 1994.
- **México Data Bank:** Indicadores económicos de México con análisis detallado.
- **Las 500 empresas más importantes de México:** Contiene información de las variables financieras, rendimientos, liquidez, etc. La información contenida va desde 1996 a 1998; entre los datos que contiene se encuentran: datos de la empresa, razón social, ejecutivos y personal total de la empresa, clase o tipo de empresa, clasificación por giro, ventas netas, etc.

- **Breviario Ambiental Mexicano:** Disco compacto que reúne todo lo relacionado con cuestiones ambientales en México así como regulaciones, decretos y normas de ecología.
- **Diario Oficial de la Federación:** Texto completo del Diario Oficial de la Federación desde 1917 a abril de 1996.
- **Aprendiendo a navegar en la Internet:** Curso practico y básico para el uso y manejo de Internet.
- **Aprendiendo Microsoft Windows 95:** Curso práctico y básico para el uso y manejo de Windows 95.
- **World Trade Almanac 1997:** Información sobre el comercio mundial país por país, datos de importaciones, exportaciones, comportamiento de las industrias y sectores.
- **OECD Economic Surveys 1996:** Estadísticas económicas de los países pertenecientes a la OECD, incluyen mapas y gráficas además de dar una explicación de las industrias y los sectores del país, brinda un estudio detallado del perfil de cada país.

4.4.9 DESCRIPCIÓN DE LAS PRINCIPALES BASES DE DATOS INSTALADAS EN DISCO DURO EN EL CEI

- **Base de Datos "Día" versión 5.0:** Diseñada como una eficiente herramienta de consulta permanente, contiene información estructurada para exportaciones e importaciones. Partiendo de las Tarifas del Impuesto General de Importación y Exportación, incluyendo sus restricciones y preferencias, los Tratados de Libre Comercio y toda la

legislación aduanera y del comercio exterior para apoyar una eficiente toma de decisiones.

- **Carpeta Electrónica del Banco de México:** Indicadores económicos del Banco de México. Actualizada mensualmente y con opción de búsqueda por palabras y por tema.
- **Banco de Datos del INEGI (BDI):** Indicadores de la economía nacional y de algunos indicadores de Estados Unidos de Norteamérica y el comportamiento a nivel estadístico de los indicadores económicos de coyuntura.
- **Tablas económicas mundiales:** Indicadores económicos del Banco Mundial por países de 1978 a 1993. Contiene información de todos los indicadores de coyuntura desglosados.
- **Tratado de Libre Comercio (TLC):** Texto integro del TLCAN.

Las bases de datos anteriores son las más importantes del Centro Electrónico de Información, los 260 títulos que figuran en el conteo de las bases de datos van desde títulos de multimedia, enciclopedias y diccionarios, hasta bases de datos especializadas en algún área en particular. En resumen se puede decir que la totalidad de las bases de datos del CEI abarcan las siguientes áreas del conocimiento con los tipos de materias que se detallan:

Ciencias Sociales y Humanidades:

Publicaciones periódicas

Diarios nacionales e internacionales

Computación:

Enciclopedias de Computación

General:

Almanaques

Enciclopedias generales y especializadas

Ingeniería:

Índices y resúmenes

Estándares

Nacional:

Estadísticas poblacionales

Legislaciones

Índices y resúmenes

Notas periodísticas

Análisis económicos

Negocios:

Compañías nacionales e internacionales

Estadísticas e información internacional

Directorios de compañías

4.5 SELECCIÓN DE LA BASE DE DATOS PARA LA APLICACIÓN DEL**CUESTIONARIO**

La base de datos que se eligió para hacer esta investigación es INFOLATINA, la razón por la cual se pensó en esta base es por que se encuentra dentro de las primeras tres

bases de datos de la Biblioteca Digital del Sistema ITESM³⁹ de mayor uso por la comunidad del Campus, posee una característica especial por ser la única base de datos que recopila fuentes de información nacionales y en idioma español, por la facilidad de acceso, ya que esta base puede ser consultada dentro y fuera del Campus a través de Internet.

Infolatina provee información sobre México y en general todo lo que se emite en las fuentes de información nacionales. Cuenta con diferentes fuentes tales como: revistas, periódicos, estudios, estadísticas, el Diario Oficial de la Federación, boletines de prensa, información de la Bolsa Mexicana de Valores, directorios especializados, agencias noticiosas internacionales, circulares, algunas leyes del Distrito Federal, discursos y comunicados de prensa oficiales o gubernamentales, estudios del ámbito económico, mercadológico, de exportación, etc. Contiene toda esta información en texto completo, está disponible para que el usuario pueda grabarla en formato ASCII o texto.

La base de datos de INFOLATINA, es una de las más importantes para la comunidad del ITESM CCM; anteriormente se tenía en suscripción únicamente para usuarios del Centro Electrónico de Información (CEI), pero, por la importancia de la información que contiene y la frecuencia de uso de toda la comunidad del Sistema, actualmente se integró INFOLATINA al proyecto de BIBLIOTECA DIGITAL, en la cual todas las Bibliotecas del Sistema ITESM tienen acceso a la misma información. Con la Biblioteca Digital se puede obtener información completa en textos e imágenes de artículos sobre diversas áreas del conocimiento como negocios, ingeniería, ciencia y

39 En la Biblioteca Digital del Sistema ITESM cualquier alumno y empleado del Tec puede tener acceso a millones de artículos de revistas y fuentes noticiosas, además de miles de libros, simposios, y más en constante actualización y sin restricciones de tiempo.

tecnología, computación, economía ciencias sociales, telecomunicaciones, educación salud y ciencias médicas y otras áreas.

Por lo anterior la base de datos INFOLATINA, se convierte en una de las principales bases de datos que apoyan la investigación al nivel del Sistema ITESM, por las razones anteriores ya explicadas INFOLATINA fue elegida como modelo para el curso de capacitación acerca del "Desarrollo de técnicas de búsqueda y recuperación de información en bases de datos del CEI". Con el objeto de que se conozca la base de datos de INFOLATINA y la manera de realizar búsquedas que describirán a continuación todas las herramientas y pantallas que brinda la base:

INFOLATINA

En esta pantalla se pueden elegir operadores o poner simplemente la búsqueda, a continuación se describirá cada concepto u operador que aparezca en la pantalla:

digital como principal

Búsqueda ?

Describa su consulta Búsqueda

Operadores:

Escoja la antigüedad deseada o el rango de tiempo.

Antigüedad Rango

Todo el tiempo / /

A / /

Máximo número de resultados por página:

Inicio Búsqueda Ayuda

A continuación se explica cada uno de los campos de la base de datos de información:

Describe su consulta: En este campo se hace la petición o búsqueda que se desee que realice la base; no importa si los caracteres indicados están en mayúsculas o minúsculas y si se incluyen acentos o no.

Antigüedad: Si se elige este campo en vez de "Rango", se tiene la opción de escoger diferentes periodos y así, dará por entendido que realizará la consulta en "Todo el tiempo", en "12 meses", en "4 meses", en "1 mes", en "1 semana" u "hoy"; o sea que, este es el periodo en el que se quiere que se realice una búsqueda.

Rango: Permite establecer un rango de fechas específicas para que la búsqueda se haga solamente dentro de ese rango de tiempo y así delimitar más los resultados.

Máximo número de resultados por página: Esto se refiere sencillamente a cuántos resultados por página se quieren, los cuales pueden ir desde 10 hasta 100. Se recomienda elegir 10 o máximo 20 resultados por página para que el proceso de recuperación de información sea lo más rápido posible.

Operadores de búsqueda: Aquí se despliegan varios operadores para delimitar la consulta. Las opciones que se pueden emplear son:

- **Sin operador.** El operador por "ausencia" es "Y". Por lo tanto, los espacios que hay entre cada palabra de una búsqueda se entenderán como "Y". Por ejemplo: elecciones estatales. El sistema buscará artículos que hablen sobre "elecciones" y sobre "estatales", pero no necesariamente juntos.
- **Operador "O".** Al utilizar el operador "O", el sistema busca artículos que contengan una u otra palabra de la búsqueda realizada. Ejemplo: elecciones


- O estatales. INFOLATINA busca artículos que hablen sobre "elecciones" o sobre "estatales", pero no necesariamente los dos términos.
- **Operador ' '.** Las comillas ' ', le indican al sistema que se está buscando artículos que coincidan exactamente con los términos deseados tal como se escriben y sin que alguna palabra esté interpuesta entre los conceptos "elecciones estatales"
 - **Operador NO.** Utilizando el operador "NO", el sistema busca artículos que no mencionen el término que se encuentra después de este operador. Elecciones estatales NO Chihuahua. El sistema busca artículos que hablen sobre "elecciones" y sobre "estatales" pero no sobre "Chihuahua".
 - **Operador *.** Este operador "*", le indica al sistema que busque artículos que contengan declinaciones de la palabra seleccionada. Ejemplo: elec*; el sistema busca : electoral, elecciones, electorales, elección, electo, etc.
 - **Operador CERCA.** El sistema busca artículos que contengan los términos deseados, pero separados hasta "cierto número de palabras" entre sí. Ejemplo: elecciones CERCA/5 estatales. INFOLATINA busca artículos que hablen sobre "elecciones" y muy cerca la palabra "estatales", tan cerca que no haya más de cinco palabras entre ambos términos.
 - **Operador \$.** Al utilizar el operador \$, el sistema busca documentos que contengan un solo carácter adicional o ninguno a la palabra seleccionada. Por ejemplo: hij\$, aquí el sistema busca: hijo, hija hijastro, etc.
 - **Operador ?.** Sustituye un carácter. Por ejemplo: ro?a. En este caso el sistema encontrará roja, roma, rosa, etc.


- **Operador mismo/**. Este operador es un operador bidireccional, ya que busca palabras dentro de "un número de párrafos adyacentes".
elecciones mismo/1 federales. El sistema encontrará aquellos documentos en los que elecciones y federales aparezcan en el mismo párrafo.
- **Operador nomismo/**. Este operador pretende que dos palabras no aparezcan dentro del mismo párrafo. Por ejemplo: elecciones no mismo/1 federales. El sistema encontrará aquellos documentos en los que elecciones y federales no aparezcan en el mismo párrafo.
- **Operador ()**. Se emplea como un delimitador en la búsqueda. Elecciones y (estados o estatales) NO (Distrito Federal o área metropolitana). El sistema buscará documentos que contengan elecciones y estados, o elecciones y estatales pero en ningún documento aparecerá Distrito Federal y área metropolitana.
- **Operador ()/f:TITULO**. Este operador le indica al sistema que encuentre la frase que se indicó entre los paréntesis en el título. Ejemplo: (elecciones estatales)/f:TITULO. Encontrará todos aquellos artículos que hablen de elecciones estatales en el título.


La pantalla que se presenta a continuación sirve para seleccionar las fuentes de información en las que se quiere que se realice la búsqueda.


digital como principal










Seleccionar fuentes de información:



Fin de sesión


Búsqueda


Ayuda


Búsqueda Internacional

 Gobierno	 Revistas	 Noticias	 Directorios	 Finanzas
 Estadísticas	 Estudios	 Internacional	 Bibliotecas	


Legal

- Circulares.
- Jurisprudencia.
- Diario Oficial de la Federación.
- Legislación del D.F.
- Legislación del Estado de México.
- Legislación Federal.

Seleccionar Fuentes de Información: Este apartado es sumamente importante, ya que se tienen muchas fuentes de información, las cuales se pueden ir seleccionando una por una o hacer la misma búsqueda seleccionando varias fuentes simultáneamente para que el sistema realice la búsqueda. Las fuentes de información con subdivisiones, aunque además cada subdivisión tiene sus fuentes específicas, son las siguientes:

- **Gobierno**

Diario Oficial de la Federación

Discursos y Comunicados de Prensa

Planes y Programas

- **Revistas**

Ecología

Fiscal - Laboral

Negocios, Economía y Finanzas

Política

Tecnología

- **Noticias**

Agencias Noticiosas

Comunicados de Prensa

Periódicos

Síntesis Informativas

- **Directorios**

Asociación Mexicana de Gas Natural, A. C.

Bancomext Trade Directory of Mexico 1997

Directorio de Emisiones - BMV

Directorio de Empresas más Importantes de México

Directorio de la Industria Química Mexicana

Directorios Internacionales - Bancomext

Orovip, Directorio del Sector Financiero de México

Orovip, Directorio del Sector Gobierno de México

- **Finanzas**

Análisis

Empresas

Valores

- **Estadísticas**

Analistas Independientes

Anuario Estadístico de la Industria Química Mexicana - 1996

Anuario Estadístico de los Estados Unidos Mexicanos 1993

Bancomext

Estadísticas de Inversión Extranjera

Estadísticas del Distrito Federal

Indice Metropolitano de Precios

Sac - Indicadores Económicos

Series de Tiempo - Banco de México

- **Estudios**

Bancomext

Comisión Nacional de Seguros y Fianzas

Consultores Internacionales

Estudios Económicos - CISLE

Instituto Mexicano de Estudios Políticos A.C.

Plano de Información Mercadológica

Reporte de Coyuntura

Select - IDC

Yes Comunicaciones - Distech Report

- **Internacional**

Catálogo de Publicaciones Bancomext 1998

OpCit-Fichas Biográficas

OpCit-Novedades Bibliográficas

- **Bibliotecas**

Catálogo de Publicaciones Bancomext 1998

OpCit-Fichas Biográficas

OpCit-Novedades Bibliográficas

Tesis de Administración - ITAM

- **Legal**

Circulares.

Jurisprudencia.

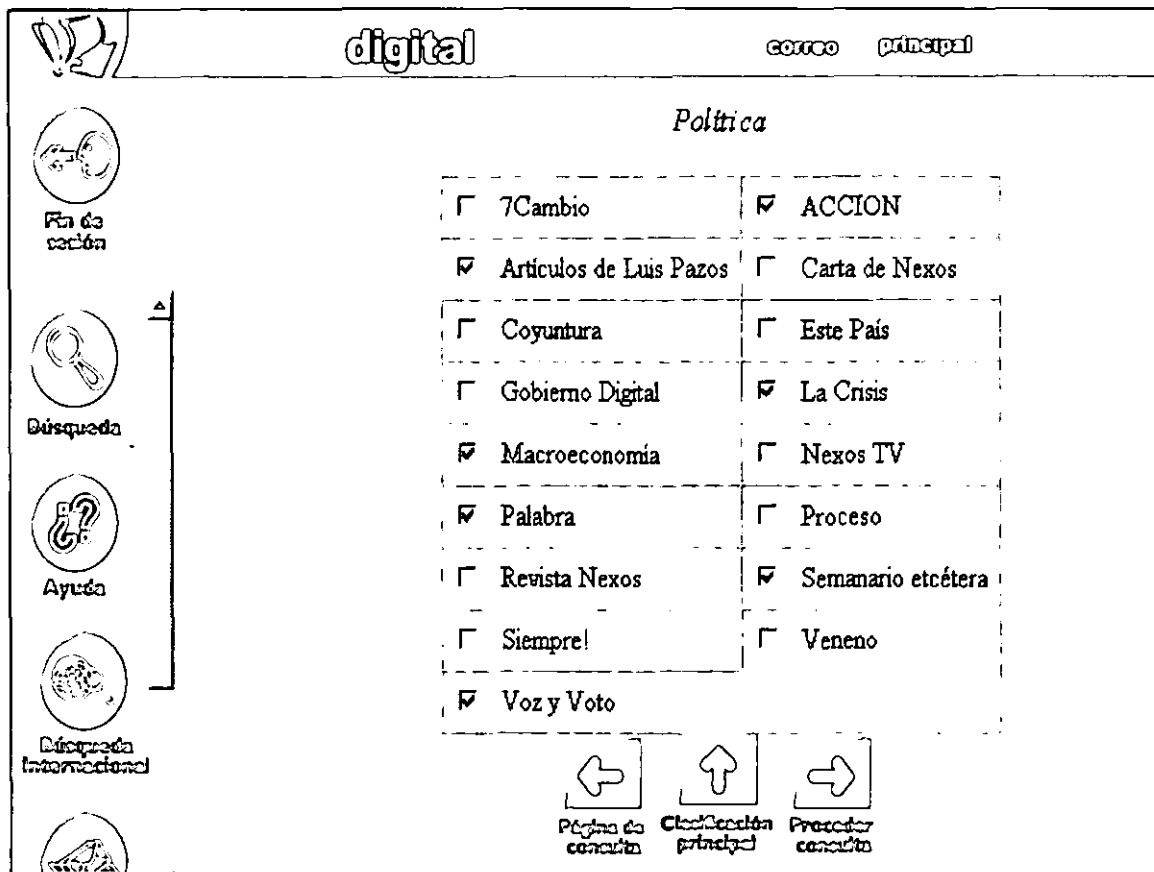
Diario Oficial de la Federación.

Legislación del D.F.

Legislación del Estado de México.

Legislación Federal.

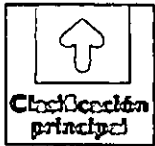
Habiendo visto cuáles son las fuentes de información disponibles en la base de datos Infolatina se pueden ir seleccionando las fuentes que se requieren dando un "click" en cada cuadro que se encuentra a la izquierda de cada fuente, y así de esa manera aparecerá la selección. A continuación en la pantalla siguiente se verá cómo aparecen las fuentes seleccionadas:



Si se desea agregar más fuentes de información a la búsqueda, se selecciona la opción "Página de consulta", para regresar a la pantalla de la elección de las fuentes con la siguiente flecha:



O si lo que se prefiere es ver en qué otro tema están divididas las revistas, se escoge "Clasificación principal", así se podrán seleccionar más fuentes de un mismo tipo, por ejemplo, de revistas.



Si ya se describió la consulta y se seleccionaron y marcaron las fuentes de información, se oprime:

Proceder consulta: Este comando se selecciona una vez que ya se hayan definido los parámetros de la búsqueda y las fuentes de información en las que se desea realizarla. De aquí en adelante se procederá a realizar la búsqueda.



HOJA DE RESULTADOS

Cuando se elige "Proceder consulta" se despliega una lista de resultados encabezados por ciertos términos:

- **VAL:** Indica la importancia del resultado de la búsqueda respecto a la consulta. La relevancia de una consulta se determina en base a cuatro criterios.

Frecuencia: El documento en el que aparezcan más veces los términos buscados será el que reciba mayor puntuación.

Densidad: El artículo o reporte con mayor densidad de aciertos recibe mayor puntaje. La densidad está dada por el número de aciertos entre el tamaño del documento.

Frecuencia inversa: Da mayor puntaje a la palabra de la búsqueda que es que es más particular.

Amplitud de aciertos: Cuanto más variados sean los aciertos en un documento, mayor será el puntaje recibido.








- **Título:** Dentro de este cuadro se desplegarán los títulos de los resultados de la consulta.
- **Fecha:** Muestra la fecha de la publicación del resultado. En el caso de legislación, la fecha corresponde a la de la última actualización que se haya hecho de una ley o un acuerdo.
- **\$ pesos:** Muestra el costo por el resultado, pero no se cobra al usuario pues existe un contrato de por medio con el proveedor en relación a un acceso ilimitado con todo el Sistema ITESM.
- **Fuente:** Muestra la fuente de información de donde se obtuvo el título, que puede ser desde un periódico hasta la información financiera de la Bolsa Mexicana de Valores.
- **PAGS:** Indica el número de paginas de que consta el documento.
- **Artículos Seleccionados:** Muestra el número de resultados que se seleccionó.
- **Cambiar la consulta:** Se emplea cuando se desea modificar la consulta.
- **Pagina siguiente:** Al oprimir este botón es posible ver la siguiente pagina de artículos encontrados.
- **Pagina anterior:** Oprimir este botón si se desea regresar a la pagina anterior de los artículos.

Hoja de resultados

Consulta: **elecciones estatales**

254 resultados. Mostrados: del 1 al 20

Cambiar
consultaPágina
siguienteSeleccione el icono  del artículo de su interés

	388	Ganar el gobierno sin perder el partido	01/04/1996	Palabra	15
	382	Las mutaciones del régimen político mexicano	01/07/1996	Palabra	21
	377	¿La brega finita?	01/07/1996	Palabra	8
	376	Equilibrio de poderes	01/07/1997	Palabra	11
	374	Tiempo de ruñanes	26/12/1998	La Crisis	8
	374	Los votantes dijeron NO a la moralina republicana	07/11/1998	La Crisis	3
	374	EU: revancha de mexicanos contra el racismo de republicanos	07/11/1998	La Crisis	5



Cuando ya se seleccionó algún resultado de las opciones que brinda Infolatina en la primera columna de la izquierda, se puede observar el texto completo del artículo; apareciendo la referencia y enseguida el cuerpo del artículo como se ve en la siguiente pantalla:



TITULO: Un país prerrevolucionario: López y Rivas (PRD)

FECHA: 26/12/1998

FUENTE: La Crisis

AUTOR: Vicente Bello

LUGAR: México, D.F.

Pese a la crisis, deseos de participar: Rodríguez Prats

Será 1999 el año "de la política de la fuerza, de la sinrazón", afirma Gilberto López y Rivas, diputado del PRD, quien añade: "La debilidad de un jefe del Ejecutivo que vive en una 'realidad virtual', la crisis económica profundizada este año, la explosividad que se agudizará por roces sociales y los millones de jóvenes que ya se dieron cuenta de que no tienen futuro ponen en el filo del incendio a un país convertido en una verdadera pradera seca".

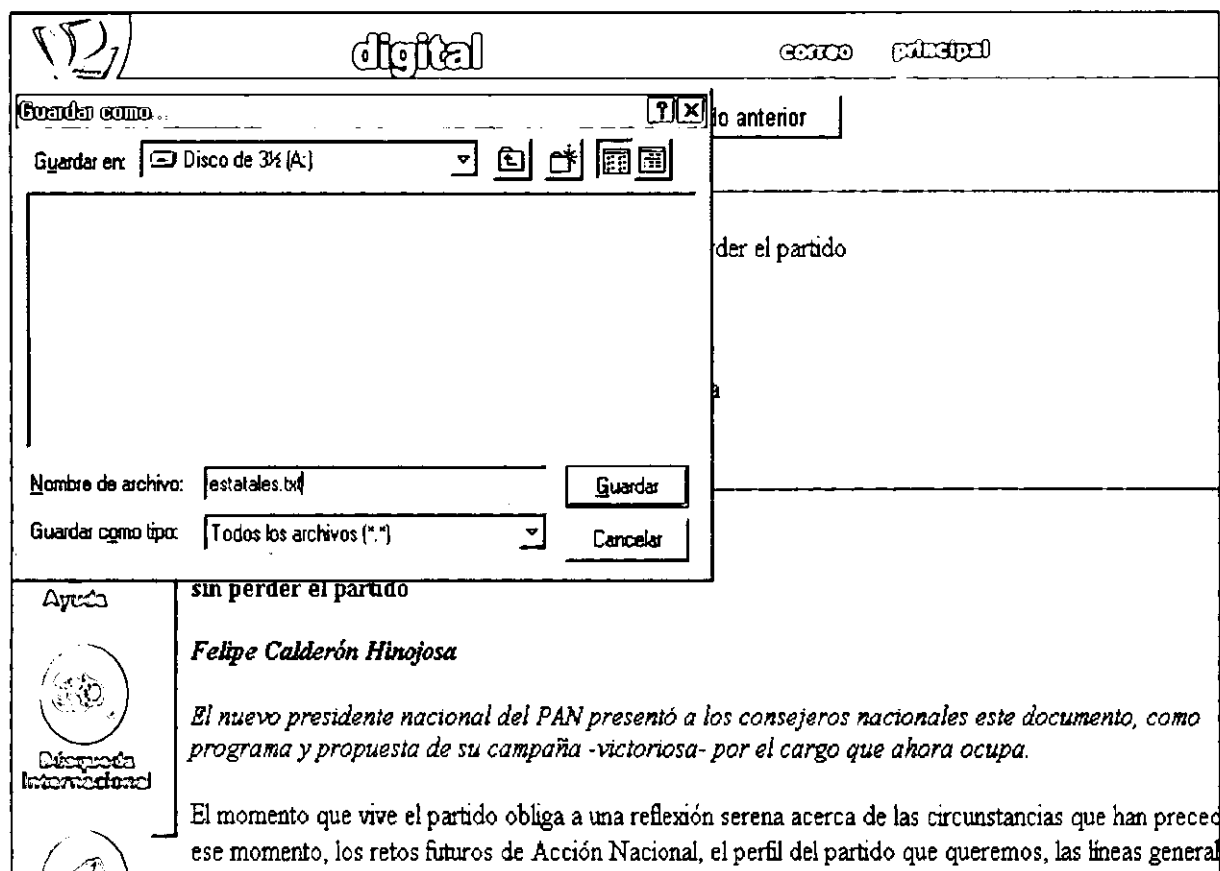
Además se pueden percibir los términos que se dieron para la búsqueda de información resaltados en color azul, por lo cual se podrá ver cuántas veces existe el término o los términos en el artículo; el siguiente ejemplo es parte del texto en donde para la búsqueda se eligió una frase exacta: "elecciones estatales", aquí se ve cómo aparece la frase resaltada en el texto:

Y en medio de todo esto, el año que viene podría ser también el de la crisis de los partidos. Crisis que se manifiesta, según diputados priistas y perredistas, en el rebasamiento de una sociedad inconforme.

Diputados panistas y priistas afirman que el año que viene será "inédito en la vida política del país con riesgo de roces sociales" y aún con la amenaza de que en las elecciones estatales éstas se estremezcan por una

GUARDAR LA INFORMACIÓN

Para guardar o grabar la información que se encuentra se realiza con la extensión .txt, por ejemplo si el nombre que se dará al archivo es de "estatales" se pondrá en el cuadro "estatales.txt" en el menú de Archivo de Internet se da "Guardar como" y luego el cuadro de guardar como archivo de texto como en la siguiente gráfica:



Este es el procedimiento que comúnmente se tiene que seguir para hacer la búsqueda y para grabar la información que se desee. Es un proceso rápido, sencillo y sobre todo, hay que indicar que la base de datos de INFOLATINA tiene la "ayuda" permanente y esto es de gran utilidad al usuario.

CAPÍTULO 5

DIAGNÓSTICO PREVIO DE LAS BÚSQUEDAS DE LA INFORMACIÓN EN BASES DE DATOS DEL CEI

En este capítulo se aborda lo relacionado al estudio que se realizó con alumnos del ITESM CCM y con la base de datos INFOLATINA, la cual anteriormente ya fue descrita con el fin de mostrar qué tipo de información se puede encontrar en ella y cuál es el procedimiento adecuado para hacer una búsqueda.

Primeramente, lo que se pretendía era analizar qué tan familiarizados están los usuarios con las bases de datos y la recuperación de información, además se investigó a través de un cuestionario, el grado de satisfacción que los usuarios obtienen al realizar sus búsquedas, se detectó si existía la intención por saber cómo hacer búsquedas de información de manera rápida, fácil y con éxito, acorde a las necesidades específicas de información de cada usuario, por lo que el curso se estructuró para ser impartido a la mitad de la comunidad seleccionada.

Por último, se hizo una evaluación de las respuestas obtenidas en el cuestionario para conocer la conformidad de los usuarios con sus procesos de búsqueda y recuperación de información en las bases de datos del CEI; en esta evaluación se tomaron en cuenta ciertos factores como son: cuáles son las razones por las que no acuden a buscar en bases de datos y cuál es el porcentaje de errores, etc. Además es importante recalcar que este cuestionario se aplicó a toda la población a investigar, sin importar si quedaron en el grupo control o en el experimental, y más adelante se realizó la selección o división de los dos grupos.

5.1 EXPERIENCIA A PRIORI

De acuerdo a la experiencia en el CEI y a las fuentes teóricas que se consultaron para poder realizar el proyecto, se puede afirmar que actualmente, a pesar de la avanzada tecnología de información para la investigación que existe, no hay un conocimiento previo que se haya impartido en las aulas acerca de las técnicas de búsqueda y recuperación de información en bases de datos. Los usuarios acuden al CEI a consultar las bases de datos por que saben que es más "rápido" buscar en formatos electrónicos que ir a la hemeroteca a consultar revista por revista o a consultar un libro, sin embargo cuando acuden a realizar sus investigaciones en las bases de datos no llegan con una idea acerca de cómo y/o en dónde buscar; al consultar las bases de datos no saben cuál es la forma más fácil de buscar o cuáles son los comandos más útiles que les podrían ayudar en su consulta y nisiquiera saben qué tipos de bases de datos existen, entre otros problemas .

En la experiencia a priori, acumulada por el trabajo cotidiano de asesorar por más de dos años a los usuarios del CEI en las búsquedas que hacen en las bases de datos, se detectó que hay una escasez de conocimiento de lo que son las bases de datos y los diferentes tipos que existen, ya que actualmente se sabe que dependiendo del formato hay muchos factores que ayudan a determinar si una búsqueda será o no exitosa. También se sabe que muchas bases de datos poseen opciones de ayuda o campos que están hechos para facilitar el proceso de búsqueda y recuperación de la información, pero en otros casos el usuario no ve más allá del cuadro de búsqueda y no explora la base con el fin de ver cual es la ayuda o los campos que lo pueden auxiliar para hacer más específica su búsqueda con relación a la información que necesita. Por esta experiencia se propone que los usuarios pueden tener una

capacitación en el uso de las bases de datos, la búsqueda y la recuperación de la información para que lleguen a tener conocimientos acerca de los tipos de bases de datos que existen, cuáles son las técnicas para obtener resultados exitosos en su investigación y cuando se dispongan a buscar en una base de datos, vean antes cómo funciona y cuáles son las opciones que pueden utilizar, los comandos más comunes y las operaciones que pueden emplear con cada base de datos logrando así que el proceso de búsqueda sea más fácil, rápido y exitoso. Por lo tanto la capacitación es una herramienta útil para lograr incrementar la cultura acerca de las técnicas de búsqueda y recuperación de información en bases de datos.

Teniendo en cuenta que se capacitarán mejor en las búsquedas, entonces, se puede decir que acudirán a consultar información en este tipo de instrumentos electrónicos en mayor número. Los resultados de las investigaciones que realicen los usuarios serán más rápidos y fáciles de obtener, por lo tanto se tendrán más usuarios e investigaciones de mayor calidad.

5.2 CUESTIONARIO

Este cuestionario se diseñó para tener una visión acerca de cómo se encuentra la situación actual de las investigaciones o consultas a través de bases de datos, qué tanto conocen o no los usuarios para buscar información en bases de datos, cuánta gente recurre a realizar sus consultas y por qué; además, se medirán cuáles son los errores más frecuentes en la búsqueda y recuperación de la información y además, como información adicional, se detectará el interés de los usuarios por un curso que trate de este tema.

Se aplicará un cuestionario a toda la población a investigar y los resultados se analizarán de acuerdo a cada grupo. Al grupo experimental se le asignará el número 1 (para diferenciar los cuestionarios) y si son del grupo de control su número es el 2. Más adelante se presentarán los resultados del cuestionario.

A continuación se presenta el instrumento que anteriormente se explicó, el cual se aplicó a los usuarios que acudieron al CEI a realizar una búsqueda.

CUESTIONARIO

Este cuestionario está diseñado para medir el nivel de satisfacción en la búsqueda y recuperación de la información en bases de datos, los resultados serán útiles para mejorar el servicio del CEI. De antemano, gracias por tu colaboración.

Contesta por favor las siguientes preguntas:

1. ¿Qué tan frecuentemente acudes al CEI?

Diariamente

Dos o tres veces a la semana

Nunca

2. ¿Por qué y acudes al CEI en busca de información?

3. ¿Has realizado alguna vez búsquedas de información en bases de datos?

SI

NO

4. ¿Qué tipo de bases de datos fueron las que consultaste?

En línea o por Internet

En CD-ROM

En disco duro

5. ¿Encontraste dificultades al llevar a cabo tu búsqueda?

Muchas

Algunas

Ninguna

6. Si encontraste dificultades describe cuáles fueron las más frecuentes

7. ¿Cómo consideras que son tus búsquedas de información?

Exitosas

Medianamente exitosas

Nada exitosas

8. ¿Te gustaría aprender acerca de las técnicas de búsqueda y recuperación de información en bases de datos?

SI

NO

5.3 EVALUACIÓN DE LA APLICACIÓN DEL CUESTIONARIO

De acuerdo con los datos obtenidos a través del cuestionario se conocieron las dificultades y los errores que los usuarios encuentran al realizar las búsquedas de información en las bases de datos que comúnmente consultan en el CEI.

Después de examinar los resultados, se procedió a diseñar el curso de capacitación a partir de las necesidades específicas que los usuarios evidenciaron a través de sus respuestas.

A continuación se da a conocer los resultados de la aplicación del cuestionario:

1. La pregunta número 1 cuestiona qué tanto acuden los usuarios al CEI, 42 usuarios, (equivalente al 70% de la población) dijeron asistir al CEI Diariamente; 15 usuarios (equivalente al 25% de la población) contestaron Tres veces a la semana y el resto, o sea, 3 personas (equivalente al 5% de la población) contestaron que Una vez solamente habían asistido al CEI a buscar información a través de medios electrónicos.
2. En la pregunta número 2, se trató de identificar las necesidades de información de los usuarios pues es abierta, se cuestionaba el por qué y para qué acudía al CEI en busca de información la mayoría de las respuestas 22 (equivalente al 37% de la población) tenían que ver con la facilidad que se suponía tendrían al buscar información en fuentes electrónicas en comparación con el hecho de buscar información en libros o revistas; 9 encuestados (equivalente al 15%) dijeron que se los habían recomendado compañeros de clase o profesores; 11 usuarios (equivalente al 18%) contestaron que buscar información a través de Internet o en bases de datos les resultaba más rápido que buscarla en libros; 8 usuarios (equivalente al 13%) dijeron que la información que buscaban para su tarea no la

habían podido encontrar en libros y por eso recurrían al CEI; 6 alumnos (equivalente al 10%) que estaba más acostumbrado a las computadoras que a la Hemeroteca o a la sección de Consulta y el resto que fueron 4 usuarios (equivalente al 7%) no respondió.

3. La pregunta número 3 cuestionaba al usuario acerca de si alguna vez había hecho búsquedas de información en bases de datos; de los 60 usuarios que se encuestaron 56 contestó que si (93%) y 4 que no (7%).
4. La número 4 preguntaba sobre el tipo de base de datos que había consultado el usuario: 30 contestaron que en línea o Internet (50%), en CD-ROM 25 (42%) y los restantes que son 5 (8%) en Disco duro.
5. La pregunta número 5 se refería a saber si el usuario encontró o no dificultades al realizar sus búsquedas de información en bases de datos, la opción de Muchas la seleccionaron 41 personas (68%), Algunas 14 (23%) y Ninguna 5 (8%).
6. La pregunta número 6 se formuló para que los usuarios que en la número 5 respondieron que sí habían encontrado dificultades al llevar a cabo la búsqueda de información, describieran brevemente cuáles fueron los problemas más frecuentes con que se encontraron, casi todos los usuarios informaron que se tardaba mucho en hacer una búsqueda de información, esto denota fácilmente que era porque no sabían cómo buscar en este tipo de fuentes de información. Los resultados fueron los siguientes: 32 usuarios (53%) no encontraba documentos importantes para el tema de su investigación y por lo tanto se tardaban mucho en el intento; 17 usuarios (28%) expresaron que no sabían cómo encontrar documentos con texto completo y 7 (12%) que obtenían demasiados documentos, y 4 personas (7%) que no sabía en qué base de datos buscar la información.

7. La siguiente pregunta es la número 7 los usuarios tenían que describir de forma breve cómo consideran que son sus búsquedas de información, en la opción de Exitosas sólo se ubicaron 7 usuarios (12%); Medianamente exitosas 16 (26%) y en Nada exitosas 37 usuarios (62%).
8. Por último, en la pregunta número 8 investigaba a cerca de la posibilidad de tomar un curso informativo de técnicas de búsqueda y recuperación de la información en bases de datos y la mayoría que fueron 58 (97%) contestaron que si y 3 usuarios (2%) dijeron que no por que no tenían tiempo.

CAPÍTULO 6

PROPUESTA PARA FACILITAR EL EMPLEO ÓPTIMO DE LAS TÉCNICAS DE BÚSQUEDA DE INFORMACIÓN EN BASES DE DATOS DEL CEI

De acuerdo a las respuestas del cuestionario y por las características de los alumnos, se consideró que era conveniente la elaboración e impartición de un curso de capacitación por ser el medio más adecuado para enseñar a los usuarios en el manejo de las bases de datos y la recuperación de la información.

En este capítulo se describirá el curso que se diseñó para capacitar a los usuarios del grupo experimental de la muestra, mientras que al grupo de control no se le impartió. Al grupo experimental se le aplicó el "Curso de capacitación sobre las técnicas de búsqueda y recuperación de información" y posteriormente se evaluó con una búsqueda de información que se les pidió realizaran a través de la base de datos INFOLATINA. Posteriormente se hizo un análisis comparativo de los dos grupos; esta evaluación se realizó midiendo sus estrategias de búsqueda y comparando los resultados, tomando en cuenta diferentes aspectos, por ejemplo: si el usuario en su consulta utilizó herramientas de búsqueda como operadores, cuántos ensayos fallidos tuvo, cuántos documentos encontró, el tiempo empleado en la búsqueda y en resumen si el usuario consideró que tuvo éxito o fracaso.

6.1. FACTORES A CONSIDERAR PARA LA ELABORACIÓN DEL CURSO

Para elaborar el curso se tomaron en cuenta diferentes factores⁴⁰ tales como:

a) Disponibilidad de tiempo y duración

El curso se llevó a cabo en cuatro sesiones de tres horas cada una, las primeras dos sesiones fueron teóricas y las otras dos fueron prácticas. Se tomaron doce horas en total para el curso.

b) Número de participantes

El curso se impartió a la mitad de la población total a investigar (esta selección se encuentra explicada ampliamente en el capítulo 6 en donde está la propuesta formal de un curso de capacitación, en un apartado dedicado a la selección de los dos grupos). A 30 participantes se les dió el curso y a los otros 30 alumnos se les aplicó directamente el examen (cuestionario de búsquedas de información) para evaluar los resultados y el procedimiento que emplearon para la búsqueda de información sin ningún curso previo y se compararán los resultados de la búsqueda.

c) Instrumentación del curso

La enseñanza impartida en un curso de capacitación acerca de las bases de datos y la búsqueda y recuperación de la información, debe tener carácter teórico-práctico, ya que se necesitan explicación, aclaración de dudas y errores que los usuarios cometen y más adelante se les proporciona una práctica para que puedan instrumentar sus conocimientos en la realidad y así adquirir habilidades y destrezas para que en cualquier lugar puedan hacer búsquedas de información exitosas.

d) Apoyo de materiales

Para la aplicación del curso se necesitarán materiales de apoyo, como acetatos, ilustraciones, pizarrón y plumones, computadoras, bases de datos, entre otros materiales que existen en un centro de información.

e) Participantes

Como se había mencionado anteriormente, el curso está dirigido a los usuarios estudiantes de nivel profesional del ITESM CCM, no importando la carrera, sólo basta que decidan tomarlo cuando se les invite, pues así demuestran interés por saber más acerca del tema.

f) Horarios

Los cursos se programaron de lunes a viernes de 10 de la mañana a 4 de la tarde para tener flexibilidad con los usuarios teniendo dos cursos por día y así los estudiantes pueden elegir el que más les convenga; de acuerdo a esto existen dos talleres de tres horas cada uno: de 10:00 a 13:00 hrs. y de 13:00 a 16:00 hrs.

g) Contenido

El curso debe ser breve y centrarse en el aspecto cotidiano y práctico, con el que comúnmente se enfrentan los usuarios al realizar una consulta en las bases de datos.

h) Mecanismos de evaluación

En un principio a los individuos que no les corresponde tomar el curso de capacitación se les pidió que hicieran una búsqueda de información y a los otros participantes, al final del curso se les pidió que realizaran la misma búsqueda, con lo que se evaluó la

estrategia de búsqueda que los participantes que tomaron el curso aplicaron; esta evaluación tomó en cuenta diferentes puntos:

- Número de ensayos fallidos
- Documentos encontrados
- Tiempo empleado
- Empleo de herramientas de búsqueda como operadores, índice, tesauros, etc.
- Si obtuvo éxito o fracasó

6.2 PLANEACIÓN DEL CURSO DE CAPACITACIÓN DE “INTRODUCCIÓN A LAS BASES DE DATOS Y LA BÚSQUEDA Y RECUPERACIÓN DE LA INFORMACIÓN EN BASES DE DATOS”.

6.2.1 ANTECEDENTES

El curso de capacitación para usuarios de bases de datos del CEI fue diseñado e impartido por la autora de esta tesis en 1998, auspiciado por el Centro Electrónico de Información del ITESM CCM, en el Distrito Federal, con una participación de 30 alumnos de nivel profesional que constituyeron el grupo control a examinar. El objetivo era brindar un conocimiento a los usuarios de las bases de datos y también acerca de las técnicas de búsqueda y recuperación de la información a emplear en ellas.

El diseño del curso de capacitación implicó la necesidad de seleccionar cuidadosamente los contenidos a abarcar. Para llevarlos a cabo se investigaron aquellas actividades que se deben tener en cuenta para realizar una búsqueda de información y cuáles eran las dificultades o dudas con las que normalmente se

enfrentan los usuarios al buscar información o al realizar una consulta en una base de datos.

Se consideró que la población sería heterogénea en cuanto a la carrera que cursaban sus integrantes ya que era homogénea en cuanto a que todos eran de nivel profesional. No hubo ningún requisito para la admisión en el curso, solamente tener interés de utilizar la base de datos y haber sido seleccionado.

6.2.2 ESTRUCTURACIÓN DEL CURSO

El curso está diseñado en cinco unidades que abarcan lo siguientes temas:

a) Bases de datos

En esta unidad se consideran diferentes aspectos como son: la definición de bases de datos y los tipos de bases de acuerdo a diferentes clasificaciones.

b) Modos de acceso a las bases de datos

Se explican los modos de acceso que existen para las bases de datos y por lo tanto, cómo se manejan de acuerdo a este criterio.

c) Recuperación de la información en bases de datos

Contiene puntos importantes como: cuáles son los comandos que pueden ayudar a realizar una búsqueda de información exitosa, los operadores universales y de manera general acerca de cómo se manejan las bases de datos de acuerdo a cada necesidad de información.

d) Estrategias de búsqueda

Son de manera breve, clara y sencilla las técnicas más utilizadas para recuperar información o encontrar los temas buscados más fácilmente.

6.2.3 OBJETIVOS GENERALES

Los objetivos generales del curso son:

1. Proporcionar al usuario los elementos esenciales para que, adquiriera los recursos que le permitan conocer los servicios que ofrecen las bases de datos, así como conocer todo lo relativo a éstas.
2. Brindar de forma clara y breve todas las herramientas que se necesitan para buscar exactamente la información que se requiere a través de las bases de datos.
3. Concientizar a los usuarios acerca de la importancia y las ventajas que tiene la consulta adecuada en una base de datos.
4. Motivar a los usuarios a que consulten más la información en medios electrónicos por la rapidez, facilidad y exactitud que aquellos pueden brindar a sus necesidades de información.

6.2.4 PROGRAMA

El siguiente es el contenido del programa sobre cuya base se desarrolló el curso:

UNIDAD 1 "LAS BASES DE DATOS"

1. Definición y tipología

1.1 Bases de datos bibliográficas

1.2 Bases de datos factuales o numéricas

1.3 Bases de datos textuales

1.4 Bases de datos mixtas

1.5 Otros criterios de clasificación

UNIDAD 2 "ACCESOS"

2. Modos de acceso a las bases de datos

2.1 Conexión en línea

2.2 Acceso local

UNIDAD 3 "LA RECUPERACIÓN DE LA INFORMACIÓN"

3. Recuperación de la información

3.1 Recuperación de información en bases de datos

3.2 Tipos de recuperación de la información en las bases de datos

3.3 Indicadores de la eficacia de la recuperación

3.4 Lenguajes informáticos, métodos y técnicas

3.5 Modelo del proceso de recuperación de la información

3.6 Operadores que auxilian en la búsqueda de información

3.6.1 Operadores boléanos y lógicos

3.6.2 Operadores de expansión

3.6.3 Operadores de intervalo

3.6.4 Operadores de proximidad o distancia

3.6.5 Operadores de cualificación

3.6.6 Operadores de encadenamiento

UNIDAD 4 "LAS ESTRATEGIAS DE LA BÚSQUEDA DE LA INFORMACIÓN"

4. Estrategias de búsqueda

4.1 Cómo determinar el tema específico a buscar

4.2 Preguntas más importantes al iniciar una investigación

4.3 Errores frecuentes en la estrategia de búsqueda

UNIDAD 5 "SELECCIÓN DE LAS BASES DE DATOS A UTILIZAR"

5. Selección de las bases de datos

5.1 Variaciones de las bases de datos

5.2 Por qué elegir entre una y otra base de datos

5.3 Sugerencias a tener en cuenta para seleccionar una base de datos

6.2.5 TÉCNICA

La metodología escogida es expositiva y demostrativa porque se consideró que es la idónea de acuerdo a la información a tratar; a ello se integró la participación activa de los capacitandos.

6.2.6 MATERIAL DE APOYO

Pizarrón.- Se emplearon pizarrones blancos con plumones en donde se hicieron diferentes cuadros y se utilizó para que se apoyaran las explicaciones del curso y se pudieran ampliar gráficamente las respuestas a las dudas y las exposiciones.

Este tipo de material permitió abordar el contenido en un lapso breve y de manera ágil, permitiendo al capacitando tomar sus notas conforme se le iba explicando el tema.

Fotocopias.- Para reforzar el proceso de enseñanza-aprendizaje y para facilitar la explicación, comprensión y desarrollo de ciertos temas, se fotocopiaron los tópicos que se iban a tratar y se repartieron entre los capacitandos en la primera sesión.

Retroproyector de acetatos.- Se emplearon dos retroproyectores de acetatos, los cuales se utilizan para que puedan ver un sólo acetato varias personas en una pared. Sirvió como apoyo en las explicaciones y en la aclaración de las dudas.

Acetatos.- Se elaboraron, aparte, acetatos que ilustran la exposición y en los cuales se fueron señalando las partes más importantes de la misma.

Computadoras .- Las computadoras se emplearon con el fin de poder dar una explicación más ilustrativa del asunto a tratar. En la parte teórica únicamente se utilizó una computadora que la autora de la tesis empleó en ciertos casos, además, en la parte práctica y en la evaluativa, que se tuvo con el grupo control se les dio una computadora a cada usuario para que pudiera hacer sus búsquedas y para que se le facilitara el aprendizaje con la práctica y así poder, además, externar alguna duda al respecto.

Bases de datos.- En general a los usuarios se les dejó elegir la base de datos que ellos quisieran para que pudieran tener más facilidad al manejar no sólo una sino varias bases de datos.

Base de datos Infolatina.- Para hacer la evaluación a los dos grupos se les pidió en distintos momentos que realizaran una búsqueda en la base de datos **Infolatina**, ya que fue la elegida para poder ver qué tan beneficioso resulta un curso introductorio a las bases de datos y la recuperación de información, para el éxito de sus consultas.

6.3 CONTENIDO DEL CURSO DE CAPACITACIÓN:

En el anexo III se presentan los puntos teóricos desarrollados en el curso que se hizo con base en las necesidades de los usuarios del ITESM CCM-CEI, además, cabe indicar que el contenido de esta tesis sirvió de base para realizar el curso de capacitación, por que, los capítulos que vienen en el principio de este trabajo, abordan de forma más amplia los temas que se trataron en el curso.

6.4 SELECCIÓN DE DOS GRUPOS: EXPERIMENTAL Y DE CONTROL

Para realizar el estudio, comparar y evaluar la eficacia del curso de introducción a las bases de datos a la búsqueda y recuperación de la información se dividió a la población total a investigar en dos grupos: el de control y el experimental.

Para elegir a los individuos que serían encuestados la única característica que se tomó en cuenta es que fueran estudiantes de nivel profesional, ya que son los que mayoritariamente utilizan esta base de datos. Para seleccionarlos se aplicó un muestreo al azar, específicamente un "Muestreo intencional". Por lo común este tipo de muestreo lo realizan expertos en el tema, la idea fundamental consiste en realizar un esfuerzo deliberado para incluir unidades representativas o importantes dentro de la muestra; es decir, se utiliza el juicio de los expertos para seleccionar unidades que sean típicas de la población o que tengan mayor importancia que otras..."⁴¹ para ello se tomó en cuenta lo siguiente:

El grupo experimental y el de control se distribuyeron de la siguiente manera:

En el programa de Microsoft Excel es un programa para computadoras, el cual se conoce mejor como "Hoja de cálculo", ya que son hojas electrónicas constituidas principalmente por columnas y renglones entrelazados (matriz) que permite introducir y almacenar datos tanto texto como números, fórmulas y resultados. Excel es un programa de cálculo que a diferencia de Lotus 1-2-3, Quattro Pro y otros es el más completo y con mayores alcances que incluye entre otras potencialidades una gran diversidad de formatos para gráficas bidimensionales y tridimensionales;

41 Guerrero G., Víctor Manuel. Estadística básica para estudiantes de economía y otras ciencias sociales. México : SEP : FCE, 1989. p.245

programación de macros, compatibilidad y capacidad para cambiar información desde otros programas para hojas de cálculo y una amplia variedad de barras de herramientas.⁴²

Con este programa se numeró en la primera columna del 1 al 60 que fue la población total a investigar; después, se seleccionó completa la siguiente columna y se utilizó el comando ALEATORIO, por lo que se rellenaron las 60 celdas de la columna 2 con números aleatorios y en la tercera columna se utilizó la función REDONDEAR CON CERO DECIMALES para aproximar estos valores entre 0 y 1, de esta manera tener sólo dos opciones, de las cuales "0 es no curso" y "1 es curso de capacitación", como se muestra en la siguiente tabla:

⁴² Candelaria Cruz, Elvio. Manual básico de Excel 5 para Windows. México: Limusa, 1997. p. 13.

POBLACIÓN TOTAL	FUNCIÓN ALEATORIA	FUNCIÓN DE REDONDEO	0=NO CURSO 1=CURSO
1	0,274518003	0	NO CURSO
2	0,887099859	1	CURSO
3	0,419040114	0	NO CURSO
4	0,604284242	1	CURSO
5	0,886866109	1	CURSO
6	0,123385489	0	NO CURSO
7	0,980209964	1	CURSO
8	0,853381215	1	CURSO
9	0,268239141	0	NO CURSO
10	0,587927579	1	CURSO
11	0,248080755	0	NO CURSO
12	0,61242584	1	CURSO
13	0,845505706	1	CURSO
14	0,12286912	0	NO CURSO
15	0,008954051	0	NO CURSO
16	0,849059536	1	CURSO
17	0,325034113	0	NO CURSO
18	0,871353986	1	CURSO
19	0,778824282	1	CURSO
20	0,044603319	0	NO CURSO
21	0,260519435	0	NO CURSO
22	0,772700734	1	CURSO
23	0,405235046	0	NO CURSO
24	0,524716444	1	CURSO
25	0,451528325	0	NO CURSO
26	0,671010887	1	CURSO
27	0,209252836	0	NO CURSO
28	0,283429282	0	NO CURSO
29	0,77030472	1	CURSO
30	0,373980783	0	NO CURSO
31	0,076597734	0	NO CURSO
32	0,477673541	0	NO CURSO
33	0,443173884	0	NO CURSO
34	0,622046336	1	CURSO
35	0,328397271	0	NO CURSO
36	0,400925629	0	NO CURSO
37	0,70192979	1	CURSO
38	0,86379195	1	CURSO
39	0,512069575	1	CURSO
40	0,246627665	0	NO CURSO
41	0,341629743	0	NO CURSO
42	0,357030111	0	NO CURSO
43	0,604048723	1	CURSO
44	0,573831094	1	CURSO
45	0,106498525	0	NO CURSO
46	0,833339386	1	CURSO
47	0,437434397	0	NO CURSO
48	0,254540323	0	NO CURSO

49	0,051839247	1	CURSO
50	0,324576112	0	NO CURSO
51	0,873320507	1	CURSO
52	0,092027895	1	CURSO
53	0,017285224	0	NO CURSO
54	0,969515046	1	CURSO
55	0,818595383	1	CURSO
56	0,636579678	1	CURSO
57	0,060346722	0	NO CURSO
58	0,876488244	1	CURSO
59	0,202367336	0	NO CURSO
60	0,660937093	1	CURSO

Por ello, si no le correspondió curso, sólo se les pidió que antes de hacer sus investigaciones nos ayudaran a efectuar una pequeña búsqueda de información en la base de datos Infolatina, (ver anexo II) y mientras realizaban su búsqueda la autora de esta tesis se apoyó en la observación directa e iba registrando los movimientos y técnicas que empleaban los usuarios para realizar la búsqueda, lo que serviría para evaluar el proceso y emitir las conclusiones.

Pero, a los usuarios que por el lugar que tenían en la tabla de asignación les correspondió tomar curso, se les informó acerca de la investigación, se les invitó a tomar un taller y se les pidió ser parte del grupo experimental, del que iba a ser capacitado en "INTRODUCCIÓN A LAS BASES DE DATOS Y A LA BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN EN BASES DE DATOS". Y a los que sí aceptaron se les pidió su nombre y teléfono para que en el momento oportuno se les recordara el curso de capacitación.

Así, se impartió el curso y al final de los dos meses el grupo que participó con la actividad final hizo la misma búsqueda que los que no se capacitaron y de igual forma a través de la observación directa se tomó nota del procedimiento para la búsqueda y recuperación de información. Por lo tanto se evaluó a los dos grupos con el mismo

instrumento, en este caso una búsqueda de información con las mismas preguntas para ambos grupos.

6.5 APLICACIÓN DEL CURSO EN EL GRUPO EXPERIMENTAL

Como se mencionó anteriormente al grupo experimental se le impartió el curso de Introducción a las bases de datos y la búsqueda y recuperación de la información, el cual ayudó a que los usuarios tuvieran mejores resultados en sus búsquedas; el temario del taller y el contenido se muestra en el anexo III, anteriormente ya se explicó qué fue lo que se tomó en cuenta para el diseño del curso.

6.5.1 RESULTADOS

Al grupo experimental se le evaluó a través de algunas búsquedas de información en la base de datos INFOLATINA que es la base que se tomó de muestra. Hay que recordar que anteriormente al grupo "0", integrado por aquellos a quienes no les correspondió tomar el curso, se les pidió de antemano su cooperación para realizar unas búsquedas de información y de esta manera se obtuvieron los resultados que más adelante se analizarán.

Para llevar a cabo esta evaluación se tomaron en cuenta diversos puntos, los que se verán con detenimiento más adelante y que sirvieron como pauta para comparar a los dos grupos: el de control y el experimental.

6.6 COMPARACIÓN DE LOS DOS GRUPOS

De los 60 encuestados, el resultado que se obtuvo al evaluarlos con la búsqueda de información que se les pidió que realizaran fue de acuerdo a 5 puntos importantes, de ahí que se examinara lo siguiente:

- a) Números de ensayos fallidos
- b) Documentos encontrados
- c) Tiempo prestado a la búsqueda de información
- d) Utilizó herramientas de búsqueda como operadores, índices, thesauro, etc.
- e) Obtuvo éxito o fracaso

Estos resultados se describirán a continuación en cuadros en los cuales se podrá apreciar mejor la diferencia entre grupos.

a) NÚMERO DE ENSAYOS FALLIDOS EN PROMEDIO CON CADA PREGUNTA

GRUPO 0 (NO CURSO)	GRUPO 1 (CURSO)
16	5
20	8
13	7
18	6
30	3
34	5
20	3
14	3
19	2
18	7
13	2
13	6
11	7
13	5
10	5
11	3
7	6
10	4
12	3
13	4
11	2
16	4
18	3
10	5
14	5
15	7
9	5
7	7
10	4
15	6

b) DOCUMENTOS ENCONTRADOS

GRUPO 1 (NO CURSO)	GRUPC 2 (CURSO)
79	5
63	7
0	7
114	8
92	10
100	10
50	10
32	8
390	9
0	7
30	13
0	10
120	7
0	6
23	5
270	4
150	3
12	2
0	5
103	4
12	7
18	8
140	5
19	5
352	4
0	2
25	3
20	5
150	4
68	7

c) TIEMPO PRESTADO A LA BÚSQUEDA DE INFORMACIÓN

GRUPO 1 (NO CURSO)	GRUPO 2 (CURSO)
30 min.	8 min.
15 min.	5 min.
44 min.	20 min.
59 min.	16 min.
25 min.	13 min.
49 min.	10 min.
32 min.	8 min.
30 min.	13 min.
22 min.	15 min.
60 min.	20 min.
32 min.	26 min.
50 min.	23 min.
58 min.	20 min.
56 min.	23 min.
36 min.	15 min.
20 min.	10 min.
25 min.	10 min.
40 min.	15 min.
15 min.	20 min.
38 min.	18 min.
23 min.	20 min.
19 min.	15 min.
33 min.	10 min.
45 min.	5 min.
30 min.	10 min.
28 min.	15 min.
33 min.	16 min.
59 min.	20 min.
45 min.	15 min.
39 min.	12 min.

d) UTILIZÓ HERRAMIENTAS DE BÚSQUEDA COMO OPERADORES, INDICE, THESAURO, ETC.

GRUPO 1 (NO CURSO)	GRUPO 2 (CURSO)
NO	SI
SI	SI
NO	SI
NO	SI
NO	SI
NO	SI
SI	SI
NO	SI
NO	SI
NO	SI
SI	SI
SI	SI
NO	SI
NO	SI
NO	SI
SI	SI
NO	SI
SI	SI
NO	SI
NO	SI
SI	SI
NO	SI
NO	SI
NO	SI
NO	SI
NO	SI
SI	SI
NO	SI
SI	SI

e) (1) ÉXITO O (2) FRACASO

GRUPO 1 (NO CURSO)	GRUPO 2 (CURSO)
2	1
2	1
2	1
2	1
2	1
2	1
1	1
2	1
2	1
2	1
2	1
2	1
1	1
1	1
2	1
2	1
2	1
2	1
2	1
1	1
2	1
2	2
2	1
2	1
2	1
1	2
1	2
2	1
2	1
1	1
2	1
1	1

6.6.1 ANÁLISIS COMPARATIVO DE LOS DOS GRUPOS

En el apartado anterior se describieron los resultados y ahora se analizarán detalladamente, pregunta por pregunta para saber qué tanto impacta o sirve tener éxito en las búsquedas de información al haber tomado el curso introductorio que nos ocupa.

Este análisis se realizó tomando en cuenta tres elementos: en la primera columna aparece el nombre del grupo en la segunda es el resultado de la sumatoria de los 30 participantes y la tercera es la división para obtener el promedio del grupo encuestado (en algunos casos de tuvo que redondear la cifra).

A) Número de ensayos fallidos en promedio con cada pregunta:

Grupo 0	440 / 30	14.66 = 15 Ensayos fallidos
Grupo 1	142 / 30	4.73 = 5 Ensayos fallidos

B) Documentos encontrados

Grupo 0	2342 / 30	81.06 = 81 Documentos
Grupo 1	190 / 30	6.33 = 6 Documentos

C) Tiempo prestado a la búsqueda de información

Grupo 0	1090 / 30	36.33 = 36 min.
Grupo 1	142	14.86 = 15 min.

D) Utilizó herramientas de búsqueda como operadores, índices, thesaurο, etc.

Grupo 0	SI=9	NO=21	SI = 100%
Grupo 1	SI= 30	NO= 0	NO = 0%

E) (1) ÉXITO O (2) FRACASO

Grupo 0	1= 7	2= 23	23% ÉXITO 77% FRACASO
Grupo 1	1= 27	2= 3	90% ÉXITO 10%FRACASO

Por lo anterior se puede decir que en los 5 puntos que se evaluaron se observó una diferencia significativa entre los dos grupos, en el Grupo 0 el 77% fracasó mientras que en el Grupo 1 sólo el 10% y el resto dijo haber tenido éxito en sus resultados; claramente se concluye que es fundamental e importante tener un conocimiento previo en el uso y manejo de las bases de datos y la búsqueda y recuperación de la información en este tipo de materiales; los usuarios que realizaron su búsqueda sin tener ningún curso previo no sólo se vieron sorprendidos y agobiados por el tiempo invertido si no por la dificultad que sintieron al ir realizando las preguntas, los documentos que ellos habían recuperado (que estos eran regularmente excesivos o en el otro extremo no encontraron ningún documento y el resultado de la búsqueda fue "0 documentos") y en general los usuarios que no habían tomado el curso introductorio tuvieron que concluir y enunciar finalmente que obtuvieron "fracaso" y no "éxito", mientras que en el grupo que sí tomaron una breve inducción dijeron sentirse diferentes y más seguros al realizar su investigación y obviamente esto se reflejó en el resultado de la búsqueda al tener el 90% del grupo éxito.

En la práctica diaria al asesorar a los usuarios en las búsquedas de información pasa

lo mismo pues pocos son los usuarios que saben en qué base de datos buscar la información que necesitan y la mayoría de las veces se pierden en otras bases o con otros temas pues no delimitan bien su tema ni tienen bien definida su estrategia de búsqueda, esto mismo pasa con las búsquedas libres en Internet, el usuario puede pasar todo el día navegando sin obtener un resultado que verdaderamente satisfaga sus criterios.

CONCLUSIONES Y RECOMENDACIONES

Con frecuencia, el objetivo de los usuarios que acuden a una biblioteca es hacer investigación y pocos son los que realmente saben cómo realizar de manera efectiva la recuperación de información a través de medios electrónicos; esto está cambiando, pues cada vez son más los usuarios que se interesan en este tipo de material dadas las ventajas que tiene en comparación con el formato impreso. Sin duda alguna, cuando los usuarios requieren información esperan que el proceso de su búsqueda sea sencillo y preciso para tener rápidamente lo que necesitan. Las investigaciones podrían ser más eficientes en la medida que el usuario conozca los recursos de información y así pueda identificar y comprender su búsqueda formulando estrategias eficaces con un auxilio mínimo del bibliotecólogo, asesor o facilitador de la información para apoyar al usuario-investigador en el uso de las bases de datos de información. Así, de manera rápida, eficaz y precisa se ahorrarían recursos económicos y humanos. A la par de esto es importante recalcar al bibliotecario o asesor la necesidad y obligación que tiene no sólo de conocer el contenido de las bases y dominarlas, sino de contribuir a que otros puedan emplearlas impartiendo cursos de capacitación con el fin de informar, sensibilizar y concientizar a los alumnos sobre el uso de las tecnologías de información. En instituciones como el ITESM que van a la vanguardia e innovación en lo relacionado al uso eficiente de la informática y las telecomunicaciones, se pretende que los alumnos sean independientes y se promuevan hacia puestos competitivos internacionalmente en su área de conocimiento. La Misión del Sistema ITESM, como anteriormente se mencionó, incluye hacer investigación y extensión relevantes para el desarrollo sostenible del

país, por ello se considera que es conveniente y factible que se instrumenten cursos de capacitación permanentes y accesibles para que los alumnos, que a la vez son los usuarios potenciales del CEI, puedan realizar adecuadamente sus investigaciones y desarrollar habilidades en el uso de las tecnologías de la información.

Evidentemente, existe la necesidad y el interés de los usuarios de una biblioteca o centro de información por conocer el uso eficiente de las tecnologías de la información que lo favorecen no sólo como estudiante sino también como futuro investigador.

Es importante que los usuarios tomen conciencia de la utilidad de conocer los avances tecnológicos, pues esto, les permitirá manejar los diversos recursos que existen en el campo de la información; de ahí la importancia de los programas de capacitación para usuarios en este ámbito.

El prediagnóstico demostró que los alumnos no conocen este tipo de fuentes de información electrónicas, lo cual representa un vacío que se pudo comprobar a través del cuestionario aplicado a toda la población investigada.

La presente investigación aportó resultados interesantes; los asesores del Centro Electrónico de Información afirman que la capacitación que reciben los alumnos puede ser una variable determinante para su autonomía en las búsquedas y para que no dependan de asesorías en sus investigaciones. Cuando un alumno sabe buscar información en bases de datos, entiende el empleo de los operadores booleanos y las estrategias básicas de recuperación, se vuelve asiduo al Centro de Información ya que acude diariamente para realizar sus investigaciones.

La aportación y recomendación principal de esta investigación es para el Sistema

ITESM, el cual debe instrumentar e incluir en sus planes y programas de estudios un curso introductorio a las tecnologías de información que incluya bases de datos, búsqueda y recuperación de información en cualquier medio electrónico, inclusive Internet. Otra de las sugerencias o recomendaciones a las autoridades educativas del ITESM, es la necesidad de que los alumnos cuenten con un Curso Sello* titulado "Introducción a las bases de datos y a la recuperación de información" o "Uso efectivo de los recursos electrónicos de la biblioteca" con la finalidad de que los alumnos del Sistema ITESM se familiaricen con estos recursos y se les facilite conocer e investigar en bases de datos de manera efectiva. Actualmente se observa que la presión que obliga a los jóvenes a prepararse y estar enterados de los avances e innovaciones tecnológicas es fuerte ya que vivimos en un mundo que cada vez exige mayor competencia y donde se vuelve prioritario el contar con habilidades en programas informáticos, paquetes de computadoras, aplicaciones tecnológicas, bases de datos de información y navegadores de Internet. Por esto es un compromiso para el Sistema ITESM brindar a su comunidad las herramientas y habilidades para formar líderes competitivos que estén a la vanguardia en la tecnología. Esto es importante si se tiene en cuenta que el Sistema TEC se interesa porque la juventud sea capaz de competir a nivel mundial con países desarrollados que se distinguen por sus avances tecnológicos. Por lo anterior, considero que es un área de oportunidad la necesidad que existe actualmente en la comunidad del Campus Ciudad de México por saber más acerca de el uso eficiente de las bases de datos. Esta investigación sustenta lo anterior de manera concreta para contribuir a la optimización de los servicios

* Curso Sello es un curso que el Sistema ITESM marca como obligatorio para todos los alumnos, el cual distingue a esta institución del resto de las instituciones afines. Actualmente en el Sistema ITESM se tienen que tomar cinco cursos sellos en cada programa académico (preparatoria, profesional y posgrado).

educativos de la institución.

También se recomienda que se hagan sondeos estadísticos de aplicación y uso eficiente de las bases de datos del CEI para conocer en qué medida se puede aumentar la colección de manera eficiente y de acuerdo a las necesidades específicas de los usuarios. Se debe realizar una evaluación de la distribución de la colección de las bases de datos del CEI para determinar las áreas menos favorecidas y en las que se requiere invertir más.

De acuerdo a la investigación también se considera importante que se realice una campaña permanente de difusión de los servicios y recursos del CEI para dar a conocer sus adquisiciones y las bases de datos útiles a las divisiones académicas, esta difusión puede ir acompañada de una reseña que sirva a los alumnos de referencia. También es relevante involucrar al docente en el uso y explotación de los recursos del CEI a través de sus tareas e investigaciones para que ellos fomenten en los alumnos la cultura de la investigación por medio de fuentes de información no impresas, mostrándoles todas las ventajas que esto tiene y también aprovechando la oportunidad para que ellos se actualicen en este ámbito.

Aplicando las recomendaciones antes mencionadas se logrará que el CEI sea un área de investigación eficiente y útil, al alcance de toda la población del Campus Ciudad de México pues se obtendrá un conocimiento profundo y expedito de las bases de datos a través de la capacitación continua, con la realización de talleres permanentes, del curso sello y de la consulta diaria, para que así todos los usuarios conozcan las diferentes opciones que brinda la biblioteca en sus fuentes electrónicas o digitales.

OBRAS CONSULTADAS

1. Bodoff, David. "A re-unification of two competing models for document retrieval". En: *Journal of the American Society for Information Science*; New York; Jan 1999.
2. Bordogna, G. "Extending Boolean information retrieval: a fuzzy model based on linguistic variables". En: *Fuzzy Systems*, 1992. IEEE International Conference, 1992.
3. Candelaria Cruz, Elvio. Manual básico de Excel 5 para Windows. México: Limusa, 1997.
4. Caridad Sebastian, Mercedes. La teledocumentación y sus aplicaciones en la información. Madrid : Forja, 1984.
5. Casa Tiraó, Beatriz. "La información en un mundo globalizado". En: *Revista Relaciones Internacionales*. n. 71 jul-sep. 1996.
6. Casa Tiraó, Beatriz. "Modelo de anteproyecto de investigación para la instalación de redes de bibliotecas en las zonas rurales mexicanas". En, IX *Jornadas mexicanas de biblioteconomía*. México : AMBAC, 1991.
7. Coll-Vinent, Roberto. Banco de datos: teoría de la teledocumentación. Barcelona: Ate, 1980.
8. De Akin, Rose. Bases de datos: uso y administración en centros de cómputo. México: Trillas, 1988.
9. Desmarais, Norman. The librarian's CD-ROM handbook. Connecticut: Meckler, 1989.
10. Díaz Guajardo, Amado R. Régimen jurídico de la capacitación y adiestramiento. México: Confederación Patronal de la República Mexicana, 1989.
11. Echeverría, L.E. "An intuitive approach for the expression of Boolean queries". En: *Visual Languages*. IEEE Workshop, 1989.
12. Espinosa Santos, Virginia. Capacitación de personal de bibliotecas. México: UNAM, FFL, 1989. (Tesis de Licenciatura).
13. Frants, Valery I. "Boolean search: Current state and perspectives". En: *Journal of the American Society for Information Science*; New York; Jan, 1999.
14. Freedman, Alan. Diccionario de computación. Madrid: McGraw-Hill, 1993.

15. Garza Mercado, Ario. Manual de técnicas de investigación para estudiantes de ciencias sociales. México: El Colegio de México, 1972.
16. Gauch, S. "Query reformulation strategies for an intelligent search intermediary". En: *Systems in Government Conference*. Proceedings of the Annual: 1989.
17. Guerrero G., Víctor Manuel. Estadística básica para estudiantes de economía y otras ciencias sociales. México : SEP : FCE, 1989.
18. Harrison, Lauren Deborah. The impact of thesauri on information retrieval (query formulation). Rutgers: New Brunswick, 1998.
19. Held, Gilbert. Diccionario de tecnología de las comunicaciones. Madrid: John Wiley, 1997.
20. Hlava, Marjorie M.K. "CD-ROM vs. Online: can both exist?" En: *Bulletin of the American Society for Information Science*. Vol. 14, no. 1. 1999.
21. Information systems and networks: Eleventh annual symposium. Sherrod, John. march 27-29, 1974. Westport: Greenwood, 1975.
22. Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Ciudad de México. Manual de biblioteca. México: ITESM CCM, 1998.
23. Instituto Tecnológico y de Estudios Superiores de Monterrey. Principios, misión, organización y estatuto general del sistema Instituto Tecnológico y de Estudios Superiores de Monterrey. México: Instituto Tecnológico y de Estudios Superiores de Monterrey, 1986.
24. Jusseaume, Leo. J. A. Procesamiento de datos; análisis de sistemas. México: Trillas, 1970.
25. Keim, Ichelle Renee. Bayesian information retrieval. University of Washington. PDH, 1997.
26. Kuhlmann, Federico. Información y telecomunicaciones. México: Fondo de Cultura Económica, 1997.
27. Ley de capacitación y adiestramiento : procedimientos instructivos y normas para registros ante la U.C.E.C.A. (Unidad Coordinadora del Empleo Capacitación y Adiestramiento). Sergio Martínez, compilador. México: Compañía General de Ediciones, 1979.
28. Lizasoain, Luis. Bases de datos en CD-ROM. Madrid : Paraninfo, 1992.

29. López, Juan Eduardo. Camino a la excelencia: la transformación del cuerpo docente del Sistema Tecnológico de Monterrey de 1989 a 1995. México: ITESM, 1995.
30. Mahalingam, K. A tool for organizing Web information.:. Computer, June, Vol. 30, 1997.
31. Martin, James. Organización de las bases de datos. México: Prentice-Hall Hispanoamericana, 1987.
32. Mendirichaga, Rodrigo. El tecnológico de Monterrey: sucesos, anécdotas, personajes. México: Castillo, 1982.
33. Reyes Ponce, Agustín. Administración de empresas, teoría y práctica. México: Limusa, 1994.
34. Shannon, Claude E. The mathematical theory of communication. Urbana: University of Illinois, 1969.
35. Sherman, Chris. The cd rom : Handbook. New york: Intertext, 1988.
36. Tsai, Alice. Sistemas de base de datos administración y uso. México: Prentice-Hall Hispanoamericana, 1973.
37. Xiangji, Huang. "Design and implementation of a Chinese full-text retrieval system based on a probabilistic model". *Computer, Communication, Control and Power Engineering*. Proceedings. Conference: 1993 2000.

LISTA DE SIGLAS

- ASTM:** Sociedad Americana de Medición de Materiales
- BANCOMEXT:** Banco Nacional de Comercio Exterior
- BDI:** Banco de Datos del INEGI
- BMV:** Bolsa Mexicana de Valores
- CCM:** Campus Ciudad de México
- CD-ROM:** Compact Disc- Read only Memory
- CEI:** Centro Electronico de Información
- CIEN:** Consulta de Información Económica Nacional
- DBMS:** Data Base Management System
- DOF:** Diario Oficial de la Federación
- EISAC:** Enseñanza e Investigación Superior, A.C.
- IEE:** Institución de Ingenieros Eléctricos
- IEEE:** Instituto de Ingeniería Eléctrica y Electrónica de Estados Unidos
- INEGI:** Instituto Nacional de Estadística, Geografía e Informática
- ISO:** Organización Internacional de Estándares
- ITESM:** El Instituto Tecnológico y de Estudios Superiores de Monterrey
- OCDE:** Organismo para la Cooperación y Desarrollo Económico
- TLC:** Tratado de Libre Comercio

ANEXO I

REGLAMENTO DEL CENTRO ELECTRÓNICO DE INFORMACIÓN (CEI)

USUARIOS

Para el presente reglamento se consideran usuarios:

1. Los alumnos de preparatoria, profesional y de graduados del ITESM, Campus Ciudad de México
2. Todos los profesores, profesionistas de apoyo y personal administrativo del ITESM Campus Ciudad de México.

Cualquiera de las personas que no esté considerada en las categorías anteriores, deberá presentar una carta institucional por escrito a la Coordinación del CEI.

REQUISITOS

- Para tener acceso al área de servicios, el usuario debe presentar cuando le sea requerida la credencial vigente que lo acredite como tal.
- El usuario deberá traer consigo un disquete de 3.5 para poder guardar la información requerida, ya que sólo en bases específicas se puede imprimir aquí, para los demás trabajos los tendrán que imprimir directamente en el CEC o en el lugar de preferencia.

RESTRICCIONES

1. Queda prohibido utilizar Internet para trabajos no académicos.
2. Introducir alimentos y/o bebidas al Centro.
3. Fumar.

4. Jugar con el equipo de cómputo.
5. Alterar el orden y la disciplina.
6. Cambiar de lugar o alterar la configuración de equipo tanto en Hardware como en Software.
7. Rayar y/o ensuciar el mobiliario, el equipo computacional o instalaciones del CEI.
8. Queda prohibida la captura de textos, para lo cual se sugiere utilizar las instalaciones del CEC.
9. En el caso de los visitantes si el Centro se encontrase totalmente ocupado se dará prioridad a la comunidad del ITESM.
10. El material del CEI únicamente será para préstamo interno.
11. Queda prohibido la consulta del e-mail, juegos o "chats".
12. En el caso de búsquedas por Internet se dará prioridad a los usuarios que necesiten consultar bases de datos del exclusivas CEI a través de Internet.

El uso inadecuado del material del CEI será responsabilidad del usuario, por lo que cualquier daño intencional, que se haga del mismo o a las instalaciones, será evaluado por un Comité encargado.

El usuario que viole cualquiera de los puntos antes mencionados, puede ser sancionado desde con un llamado de atención, hasta con la suspensión definitiva de los servicios del CEI.

Cualquier otro punto no previsto anteriormente quedará a juicio de las personas responsables del área.

**ANEXO II (PREGUNTAS PARA EVALUAR EL DESARROLLO DE LAS
BÚSQUEDAS)
CUESTIONARIO**

¿Cuál es la empresa mexicana que se encuentra en el primer lugar en ventas en México?

¿De las empresas que cotizan en la Bolsa Mexicana de Valores, cual es la que tuvo el las acciones más altas en el primer trimestre de 1998?

¿Cuales fueron los últimos acuerdos de las cuotas en la UNAM?

¿Cuántos artículos has recuperado de la pregunta anterior y en cuántos viene exactamente la información que necesitas?

ANEXO III (CONTENIDO DEL CURSO DE CAPACITACIÓN)

A continuación vienen el contenido del curso de capacitación que se impartió al grupo experimental.

"INTRODUCCIÓN A LAS BASES DE DATOS Y A LA BÚSQUEDA Y RECUPERACIÓN DE LA INFORMACIÓN EN BASES DE DATOS"

1. DEFINICIÓN DE BASE DE DATOS Y TIPOLOGÍA

Una base de datos es cualquier conjunto de información almacenada en cualquier tipo de soporte de forma que sea "legible" o accesible y amigable por un sistema informático.

Existen diferentes tipos o clasificaciones de bases de datos, las cuales pueden ser las siguientes:

- **Bases de datos referenciales: bases de datos bibliográficas. Directorios.**
- **Bases de datos fuentes: bancos de datos factuales o numéricos**
- **Bases de datos textuales.**
- **Bases de datos mixtas.**

A continuación se explicará brevemente cada una de ellas.

1.1 BASES DE DATOS BIBLIOGRÁFICAS

Se trata de uno de los tipos de datos más frecuentes. Las fuentes originales suelen ser elementos impresos (artículos de revistas habitualmente), aunque también los documentos primarios pueden estar en otro tipo de soportes (material audiovisual, programas informáticos, etc.)

Lo importante es que del documento original, en la base de datos se almacena sólo

información de referencia. Por ejemplo, para el caso más usual de artículos de revistas científicas, la información almacenada suele responder al siguiente esquema:

- Autor o autores del artículo.
- Título del mismo.
- Nombre y datos (fecha, número, páginas) de la revista donde ha sido publicado.
- Palabras-clave o descriptores que pretenden describir el contenido del artículo.
- Resumen del mismo.

En este caso la información se estructura de tal manera que cada uno de estos apartados constituye un *campo*, y el conjunto de todos ellos configura el *registro* relativo al documento en cuestión.

A su vez, el conjunto de todos los registros conforma la base de datos.

1.2 DIRECTORIOS

En este caso la información almacenada hace referencia habitualmente a datos relativos a personas o instituciones. Los casos más frecuentes son los directorios de empresas, organizaciones e instituciones públicas normalmente agrupados por áreas de actividad.

1.3 BASES DE DATOS FACTUALES O NUMÉRICOS

Aunque se emplean indistintamente ambos calificativos, lo importante es señalar que en este caso la base de datos contiene información numérica de diversa índole y organizada según diferentes criterios.

1.4 BASES DE DATOS TEXTUALES

También se denominan de texto completo, integral o, directamente en inglés, *texto*. En

este caso, el documento original completo se incorpora a la base de datos, de tal manera que cada registro está formado no sólo, por la referencia bibliográfica del documento original, sino por el texto íntegro de dicho documento.

Los ejemplos más frecuentes de este tipo de bases de datos se encontrará en aquellos campos donde al usuario no le basta con la información referencial suministrada por las bases de datos bibliográficas y, por el contrario, requiere todo el texto del documento original. Hay más áreas de aplicación como luego se abordará, pero ahora se citará este caso, el jurídico-legal.

1.5 BASES DE DATOS MIXTAS

Por último, existen casos de bases que mezclan información de ambos tipos. La distinción entre los tipos de bases de datos establecidos no se basa en un criterio totalitarista de "todo o nada" sino en la pregunta de ¿qué tipo de información es la que mayoritariamente incorporan?, ya sea documental o textual, o numérica.

Es por esto por lo que al hablar de las bases de datos textuales decía que la información en ellas contenida no es o no era sólo de tipo referencial.

1.6 OTROS CRITERIOS DE CLASIFICACION

Las tipologías pueden establecerse atendiendo a otro tipo de criterios.

Por ejemplo, En el AREA TEMÁTICA a que hagan referencia los datos en ellas obtenidos, se puede hablar de bases de datos:

- científicas
- tecnológicas
- sociales
- económicas
- jurídico-legales

- bio-médicas
- económico-financieras

Otro posible criterio es el del AMBITO TERRITORIAL de las mismas:

- internacionales
- nacionales
- regionales
- locales

SU DIFUSIÓN:

- restringida
- públicas

DE SU ORIGEN

- *institucional*
- públicas
- privada

2. MODOS DE ACCESO

Aparte de saber los diferentes tipos de bases de datos que existen y las clasificaciones que se pueden hacer hay que señalar ahora que también hay diferencias en los accesos ya que no todas las bases de datos tienen la misma estructura, la forma de acceder las bases de datos son las siguientes:

2.1 CONEXIÓN "EN LÍNEA"

La denominación teledocumentación corresponde al conjunto de técnicas provenientes de la documentación científica, la informática y las telecomunicaciones

que permiten el acceso y recuperación a distancia de la información.

Línea de telecomunicaciones es todo medio físico, por ejemplo un hilo conductor o un haz de microondas, que se emplea para transmitir datos. Sinónimo de línea de transmisión.

La palabra "telecomunicaciones" es un término relativo a la emisión o recepción de señales, imágenes, sonidos o información por soportes de hilos metálicos, radio, fibra óptica o enlaces de infrarrojo.

El proceso de conexión en línea tiene un esquema simplificado es el de una computadora central en el que se hallan almacenadas las bases de datos que son accesibles mediante terminales remotas.

2.2 ACCESO LOCAL

La posibilidad de acceder "fuera de línea" (*off-line*) a las bases de datos viene dada por la capacidad de almacenamiento de los soportes ópticos y por el incremento de potencia y velocidad de proceso de las computadoras personales.

CD-ROM es un acrónimo cuyo desarrollo en inglés es *Compact Disc- Read only Memory* y que puede traducirse como "Disco Compacto- Memoria únicamente de lectura". *Read Only Memory* es un término genérico usado para designar cualquier dispositivo informático de almacenamiento de datos permanente y no modificable.

A veces se hace excesivo hincapié en la enorme capacidad del CD- ROM olvidando que éste es un circuito periférico más de la computadora, y que es éste último el que debe localizar y procesar tal masa de información.

CD-ROM parece especialmente adecuado para bases de datos cuyo tamaño oscile entre los 20 y los 500 Mb. Bases de datos más pequeñas pueden ser instaladas sin problemas en un disco duro de capacidad creciente.

Con respecto a las grandes bases de datos, el límite viene impuesto por el del tamaño del CD-ROM. Como son también bastantes las que exceden de 500 Mb., lo que se acostumbra a hacer es distribuirlas en varios discos.

3. RECUPERACIÓN DE LA INFORMACIÓN

Con el término recuperación de información (que en la literatura anglosajona aparece como *information retrieval*) se designan todas las técnicas relativas al acceso y posterior selección de la información almacenada en un sistema informático.

La recuperación de la información es un proceso que realiza el usuario con las bases de datos, el cual está orientado a identificar aquellos registros, en este caso “documentos”

que para el usuario sean relevantes.

3.1 RECUPERACIÓN DE LA INFORMACIÓN EN BASES DE DATOS

Cuando un usuario decide utilizar este tipo de recursos, lo primero que debe hacer es formular con la terminología propia la demanda de información que requiere.

Esta formulación debe incluir conceptos o ideas básicas, así como las relaciones que entre aquellos se establecen, expresados en lenguaje natural.

La segunda tarea a indagar es, consecuentemente con lo anterior, seleccionar la o las bases de datos en la que previsiblemente se encuentra almacenada la información relativa al tema que se investiga.

Dicha selección suele efectuarse tomando en consideración criterios como el tipo de documentos almacenados, la procedencia de los mismos, la cobertura temporal, etc.

3.2 TIPOS DE RECUPERACIÓN DE LA INFORMACIÓN EN BASES DE DATOS

Existen dos tipos de recuperación de la información, los cuales son: la "recuperación factual" y "recuperación documental".

En el primer caso, se trata de buscar una información concreta habitualmente de tipo numérico. Por ejemplo, producción de trigo en México durante el año 1990.

Se trata de consultas de tipo determinista que habitualmente se efectúan sobre bancos de datos numéricos.

Por el contrario, la recuperación documental es de tipo probabilístico y se lleva a cabo normalmente sobre bases de datos. En el primer caso, la respuesta del sistema sería una cantidad. Por el contrario, en el segundo, el sistema proporcionaría una lista de documentos.

3.3 INDICADORES DE LA EFICACIA DE LA RECUPERACIÓN

Independientemente de que sea una cuestión subjetiva la evaluación que el usuario realiza de lo ajustado o no del conjunto de documentos recuperados a sus necesidades de información, cuando se habla de eficacia en este contexto es preciso abordar dos cuestiones: **la precisión y la exhaustividad**, o si se prefiere, desde el punto de vista negativo, **ruido y silencio**.

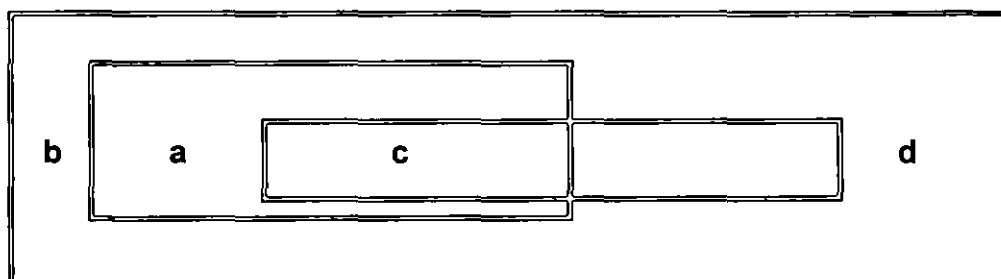
El ruido está formado por las referencias no útiles obtenidas en una búsqueda. La tasa de ruido es el porcentaje de referencias no útiles sobre el total de las obtenidas.

A la noción de *ruido* se opone la de *silencio*, que se define como el número de documentos pertinentes existentes en el fondo documental y que no han sido seleccionados en la búsqueda.

La tasa de precisión es la inversa a la del ruido y es la proporción de referencias pertinentes sobre el conjunto recuperado.

La tasa de *exhaustividad* es el porcentaje de las referencias pertinentes obtenidas en la consulta sobre el total de los pertinentes existentes en el fondo documental.

El conjunto de referencias que constituye el fondo documental de un sistema lo podemos representar según el siguiente modelo:



a = documentos pertinentes recuperados

b = documentos pertinentes no recuperados (silencio)

c = documentos no pertinentes recuperados (ruido)

d = documentos no pertinentes no recuperados

a + b = documentos pertinentes existentes

a + c = documentos recuperados

a + b + c + d = total de documentos en el fondo

$$\text{Tasa de exhaustividad} = \frac{a}{a + b}$$

$$\text{Tasa de silencio} = \frac{b}{a + b}$$

$$\text{Tasa de pertinencia} = \frac{a}{a + c}$$

$$\text{Tasa de ruido} = \frac{c}{a + c}$$

3.3 LENGUAJES INFORMÁTICOS, MÉTODOS Y TÉCNICAS

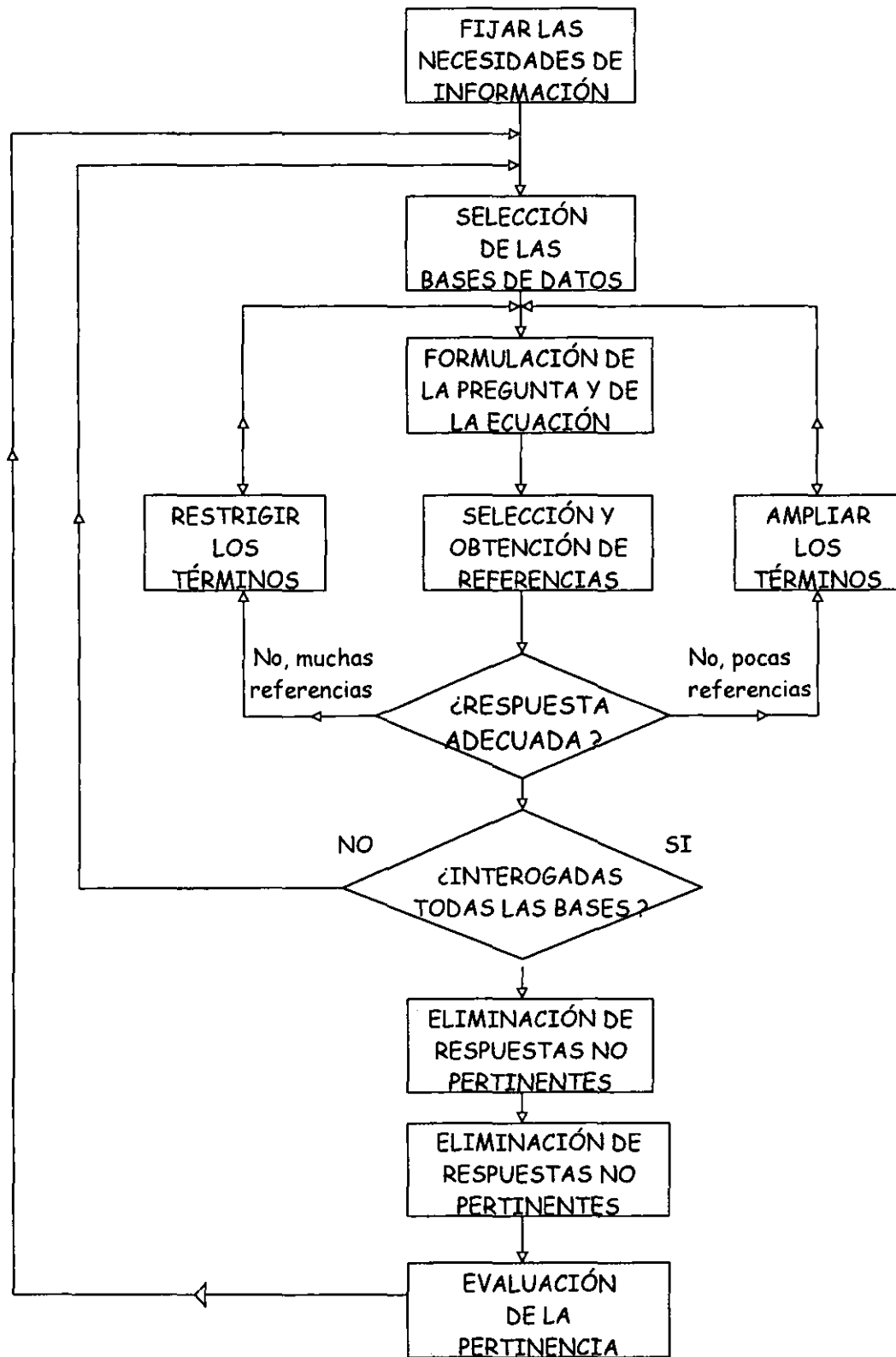
El procedimiento más simple para llevar a cabo la interrogación al sistema consiste en teclear un término junto con la instrucción adecuada del lenguaje informático que se emplee. Independientemente del tipo de lenguaje documental que se emplee, la interrogación a la base puede efectuarse de un modo más preciso si en vez de realizar este tipo de consultas simples, la interrogación se hace más compleja. En definitiva, si se formula una **ecuación de búsqueda**.

ECUACIÓN = TÉRMINOS + OPERADORES

Los términos de búsqueda pueden ser "palabras clave" o descriptores en función del tipo documental que se emplee, libre o controlado, tipo thesaurus .

La consulta a una base de datos, consiste básicamente en seleccionar del fondo documental existente en el sistema un subconjunto de referencias que reúnan unas determinadas condiciones formuladas por el usuario.

3.5 MODELO DEL PROCESO DE LA RECUPERACIÓN DE LA INFORMACIÓN



3.6 OPERADORES QUE AUXILIAN EN LA BÚSQUEDA DE LA INFORMACIÓN

Se trata del tipo de operadores más universalmente empleados, y es posible contar con ellos en todas las bases de datos y en los distintos lenguajes de recuperación.

3.6.1 OPERADORES BOOLEANOS O LÓGICOS

Le indican a la computadora la relación lógica existente entre los términos buscados a nivel registro. Esta lógica permite realizar tres tipos de enlace descriptores que se encuentran en los términos de búsqueda.

Las principales relaciones lógicas usadas son:

Unión de conjuntos (disyunción, "o" inclusivo). Indica que las palabras son sinónimos o términos relacionados. Al menos un término deberá estar presente en los registros recuperados. Utilizar el operador "O" permite enlazar dos descriptores, debiendo hallarse uno u otro, o ambos en la indización del documento para que éste sea pertinente. Es usado para incluir sinónimos. Este conector es usado para ampliar los términos de búsqueda, como también elimina duplicación de registros.

Intersección de conjuntos (conjunción, "y"). Se usa para combinar conceptos, en donde todos deberán estar presentes en el mismo registro pero no precisamente en el mismo campo. Utilizar el operador "Y" permite enlazar dos descriptores que deben encontrarse en la indización del mismo documento para que éste sea considerado pertinente. Concretamente, ambos términos deben encontrarse en cada registro.

Este operador busca por artículos que contengan todas las palabras claves seleccionadas. Este operador limita la búsqueda.

Exclusión de conjuntos (negación, "no"). Se utiliza para eliminar conceptos.

Cada conjunto está definido por una característica de análisis. El utilizar el operador "NO"; permite unir dos descriptores de los cuales el primero debe estar presente y el segundo, ajuste en la indización de un documento para que éste sea pertinente. Elimina doble significados de palabras y también limita la búsqueda.

Como resumen final, el siguiente cuadro reúne las diferentes denominaciones y operadores más usuales:

INTERSECCION, conjunción, producto : *, ET, AND, Y

UNION, disyunción, suma : +, OU, PR, O

EXCLUSION, negación, resta : -, SAUF, NOT, NO

3.6.2 OPERADORES DE EXPANSIÓN

Son aquellos que sirven para incrementar el campo de acción de la búsqueda a más de un término mediante la sustitución de caracteres en el mismo. Entre estos operadores de expansión suele distinguirse entre operadores de *enmascaramiento* y operadores de *truncamiento*.

Los operadores de enmascaramiento sirven para reemplazar un solo carácter del término de búsqueda. Por ejemplo:

MICRO?COMPUTADORA

En este caso, el carácter "?" cumple la función de "comodín" de forma que "vale" por cualquier carácter (incluida la secuencia vacía), pero uno solamente.

Ejemplo: MICROCOMPUTADORA Y MICRO-COMPUTADORA.

Los operadores de truncamiento se emplean habitualmente con las raíces de los términos. El símbolo más frecuente suele ser el asterisco *. Por ejemplo, PROCES*. Este término nos serviría para recuperar palabras como PROCESAMIENTO, PROCESO, etc.

3.6.3 OPERADORES DE INTERVALO

También se le llaman operadores relacionales, sirven para construir intervalos numéricos. Los más habituales son:

IGUAL	=
DISTINTO	<>
MAYOR QUE	>
MENOR QUE	<
MAYOR O IGUAL	>=
MENOR O IGUAL	<=

Su uso es frecuente en el campo de la fecha o con otros campos de tipo numérico o temporal.

Por ejemplo, AÑO >= 1989. En este caso, se obtendrían las referencias de los documentos publicados desde 1989, éste incluido, hasta la actualidad.

3.6.4 OPERADORES DE PROXIMIDAD O DISTANCIA

Se emplean sobre todo, en los campos de texto libre o en las bases de datos en texto completo. Sirven para definir la distancia máxima admisible que debe separar a determinados términos. Habitualmente adoptan la forma de un número que indica el máximo número de términos intermedios.

3.6.5 OPERADORES DE CUALIFICACIÓN

Aquí se trata de la posibilidad que brindan la mayoría de los lenguajes de recuperación de especificar el campo en el que debe verificarse la condición expresado mediante el empleo del resto de los operadores.

Hasta ahora, y dado que no se había especificado lo contrario, las búsquedas serían realizadas sobre todo el texto de los registros de la base correspondiente.

Así, el ejemplo de la intersección (INFORMÁTICA Y DOCUMENTACIÓN), daría como resultado que se obtuviesen como respuesta las referencias de todos los documentos en los que aparezcan simultáneamente estos dos términos, pero independientemente del lugar de registro en que se encuentren.

Así, podría haberse dado el caso de que "INFORMÁTICA", apareciese en el título y "DOCUMENTACIÓN" en el resumen.

Si se quiere especificar más y obtener una respuesta más precisa, se podría cualificar la búsqueda poniendo como condición que estos dos términos aparezcan en el título o en los descriptores o temas.

(INFORMATICA Y DOCUMENTACIÓN) EN EL TÍTULO.

(INFORMATICA Y DOCUMENTACIÓN) EN LOS TEMAS.

De esta manera se formula una expresión más compleja mediante el empleo simultáneo de varios operadores.

Esta posibilidad de definir el campo de búsqueda aumenta la precisión de las búsquedas y además disminuye el tiempo de respuesta.

3.6.6 OPERADORES DE ENCADENAMIENTO

Por último, la mayoría de los lenguajes de recuperación permiten emplear como elementos una ecuación de búsqueda los resultados de una o unas precedentes. De esta manera se pueden ir efectuando búsquedas simples y, en función de los resultados de los mismas, combinarlos para obtener una respuesta final satisfactoria.

4. ESTRATEGIAS DE BÚSQUEDA

Las estrategias de búsqueda son el conjunto de técnicas que sirven para encontrar la información que el usuario necesita, existen muchas y varían dependiendo la base, la información, las fuentes, etcétera, pero en general podemos señalar como importantes las siguientes:

4.1 ¿CÓMO DETERMINAR EL TEMA ESPECÍFICO A BUSCAR?

El éxito de la búsqueda dependerá de dos factores:

- El cuidado que se tenga en seleccionar y escribir las palabras claves o las frases descriptoras.
- De la elección de la base o bases a utilizar.

4.2 ¿CUÁLES SON LAS PREGUNTAS MÁS IMPORTANTES AL INICIAR UNA INVESTIGACIÓN?

1.- ¿Cuál es el tema principal de la búsqueda?

Esto se resuelve pensando ¿cómo se puede describir el tema en una sencilla oración? o ¿cuál sería el título del artículo perfecto?

2.- ¿Qué palabras deberá tener un artículo para que se pueda considerar que es relevante con respecto al tema principal?

- 3.- ¿El tema de búsqueda amerita usar vocabulario científico o técnico, requiere del uso sinónimo?
- 4.- ¿Todas las palabras requieren de estar en el registro localizado para considerarlo relevante?
- 5.- ¿La información deberá estar limitada por alguna característica importante como área geográfica, población, industria o actividad económica?
- 6.- ¿Existe algún término que no desee que aparezca en la búsqueda?
- 7.- ¿Existe alguna limitación respecto al idioma en que se encuentre la información?
- 8.- ¿Es necesario limitar la información a una fecha o rango de fechas?
- 9.- ¿Requiere hacer una búsqueda intensiva en todas las bases de datos del tema, o limitada a las bases de datos más comunes?

4.3 ERRORES FRECUENTES EN LA ESTRATEGIA DE BÚSQUEDA

Cuando el resultado de la búsqueda es cero, o los registros no son relevantes, los errores pueden ser:

- Una base de datos no adecuada
- Demasiados conceptos especificados
- Se especifica demasiado algún concepto
- No se emplea el truncamiento
- No se incluyen sinónimos suficientes
- Operadores de proximidad demasiado específicos
- Uso de AND cuando debe de usarse OR
- Uso incorrecto del NOT
- Términos ambiguos o no específicos

- Falta de conceptos
- Uso de OR cuando debe usarse el AND
- Truncamiento demasiado corto
- Falta de uso de sufijos (campos delimitados)

5. SELECCIÓN DE LAS BASE DE DATOS UTILIZAR

La elección de la o las bases de datos que se requieren para hacer una investigación es un tema muy importante ya que se pueden tener técnicas correctas y efectivas de búsqueda pero si no se tienen la o las bases de datos adecuadas, entonces no se obtendrán los resultados deseados; en lo siguiente se tendrán ciertas cuestiones que se tienen que considerar cuando se tiene que seleccionar la o las bases de datos.

5.1 VARIACIONES DE LAS BASES DE DATOS

- **Contenido:** negocios, legal, medicina, tecnología, etc...
- **Tipo:** citas bibliográficas, abstracts, textos completos, directorios.
- **Presentación:** textos, imágenes, tablas estadísticas.
- **Nivel técnico de los datos:** reportes de investigación, artículos de periódicos.
- **Tipo de datos:** actualizados, publicados, no publicados.

5.2 POR QUÉ ELEGIR ENTRE UNA Y OTRA BASE DE DATOS

- a) La necesidad de elegir depende del tipo de datos que el usuario requiera: bibliográficas, numéricos, directorios, textos completos, etc.

- b) es necesario verificar si el contenido de las bases de datos es de un tipo específico: histórico, si se actualiza cada determinado tiempo, temas principales, proveedor, etc.
- c) La cobertura de la información, por ejemplo si es nacional o internacional.

5.3 SUGERENCIAS A TENER EN CUENTA PARA SELECCIONAR UNA BASE DE DATOS

- Consultar primero en las bases de datos mayores y más apropiadas
- Las bases de datos menores y más especializadas ofrecen ventajas cuando se necesitan unas cuantas referencias muy pertinentes más que una investigación exhaustiva.
- El valor de una base de datos depende en gran parte de la literatura que contiene. Se recomienda analizar las fuentes de la misma.

Es importante contar con todas las herramientas auxiliares de búsqueda posibles.