

LAT 395

*J*

BIBLIOTECA



CENTRO UNIVERSITARIO  
DE INVESTIGACIONES  
BIBLIOTECOLÓGICAS

PROYECTO DE UN PROGRAMA DE MANTENIMIENTO Y  
EXPLOTACION DE LA BASE DE DATOS LIBRUNAM.

JOVV VALDESPINO VAZQUEZ

Julio de 1982.

INFOBILA

1304



<u>CONTENIDO:</u>	Página.
Introducción.	1
Generalidades.	3
Característica de la base de datos del Sistema Librunam.	5
Archivo Central.	5
Archivos Directorios.	6
Directorio Numérico del archivo Central.	8
Directorio de llaves.	11
Directorio de llaves únicas.	13
Directorio de listas de llaves.	16
Archivos Auxiliares.	16
Archivos Temporales.	20
Estrategia para el mantenimiento y explotación de la base de datos librunam.	21
Problemática.	21
Soluciones propuestas.	22
Necesidades de Software y Hardware presentes y a un plazo de diez años.	25
Software. Necesidades presentes.	25
Software. Necesidades a diez años.	26
Hardware. Necesidades presentes y a diez años.	27
Recursos de Equipo.	29
Recursos Humanos.	29
Bibliografía.	30

## INTRODUCCION

Los objetivos fundamentales de la Universidad Nacional Autónoma de México consisten en respaldar y fomentar la enseñanza, la investigación y proporcionar los medios para transmitir y difundir la cultura. Para la realización de éstos, cuenta entre otras herramientas, con un sistema bibliotecario integrado por bibliotecas ubicadas en Escuelas de Educación Media Superior, Escuelas Profesionales, Facultades, Centros e Institutos que representan en conjunto la fuente documental más completa de México.

El sistema antes mencionado, está coordinado y dirigido técnicamente por la Dirección General de Bibliotecas.

Dentro de las funciones de la D.G.B., como coordinadora en relación al material bibliográfico, sobresalen:

- Control de nuevas adquisiciones.
- Control presupuestal de las bibliotecas.
- Uniformidad en los procesos técnicos.

- Catalogación y clasificación de libros.
- Elaboración y duplicación de tarjetas.

Dado que el material bibliográfico que se tiene que someter a estos procesos se incrementa en forma considerable, lo cual hace imposible llevarlos a cabo en forma manual y da la pauta para la creación del sistema automatizado LIBRUNAM que apoya y agiliza estos procesos.

## GENERALIDADES

LIBRUNAM es un sistema integral de proceso, almacenamiento y recuperación de información bibliográfica, cuyos objetivos principales son:

- 1.- Apoyar al Departamento Técnico en los procesos técnicos de:
  - a) Catalogación y clasificación.
  - b) Elaboración de juegos de tarjetas.
  - c) Distribución de libros.
  - d) Control de procesos menores.
- 2.- Sirve al usuario de las bibliotecas en sus procesos de consulta, investigación, elaboración de bibliografías y localización de material.
- 3.- Control de calidad en los catálogos bibliográficos de las bibliotecas departamentales.

Además, es un Banco de Datos sencillo, accesible a cualquier usuario, seguro, confiable y compatible con la mayor parte de catálogos automatizados del mundo.

Actualmente la información contenida en el Banco de Datos - LIBRUNAM, es aproximadamente de 250,000 títulos.

El Banco de Datos y los módulos programados que constituyen el sistema, están almacenados en la computadora Burroughs - 6700, propiedad de la Universidad Nacional Autónoma de México.

## CARACTERISTICAS DE LA BASE DE DATOS LIBRUNAM.

La Base de Datos del sistema LIBRUNAM está constituida por -  
archivo central que contiene las fichas bibliográficas; cua-  
tro archivos-directorios, que accesan el archivo central; ar-  
chivos auxiliares y archivos temporales de trabajo en donde  
se guardan estadísticas, información histórica, etc.

Todos los archivos usan el registro cero como registro de --  
control con excepción de los archivos auxiliares y los tempo-  
rales.

Los controles que existen varían dependiendo del tipo de ar-  
chivo e incluyen: número de registro presente, tamaño de -  
los datos, apuntador al próximo lugar disponible, fecha de -  
la última versión y descripción de versión.

Para que todos los archivos sean compatibles entre sí y no -  
se pierda información.

### Archivo Central.

Este archivo es de longitud fija y la información contenida

esta alojada en forma dinámica, es decir, permite alojar cadenas alfanuméricas de longitud variable, nodos mixtos de información fija y variable como son las fichas bibliográficas.

La información se aloja con técnicas de manejo dinámico utilizando únicamente el número exacto de palabras requeridas.

Además, cuenta con ligas que pueden continuar el nodo de información en áreas disponibles en caso de ser necesario.

El acceso a este archivo es a través de un directorio que contiene la posición exacta del nodo con la información requerida.

Cada uno de los registros físicos es de 240 palabras, se escogió esta alternativa, debido a las características de la computadora B6700, que indican que es un múltiplo de palabras adecuado para un manejo óptimo de recursos ya que las áreas con las que trabaja la B6700 son de treinta palabras.

#### Archivos directorios.

Accesar un archivo central de información en forma directa es

muy importante para un sistema de recuperación, por lo que es necesario contar con directorios que faciliten la localización de la información deseada en forma directa a partir de diferentes llaves; de otra forma, se debería recorrer todo el archivo cada vez que se desee hacer una búsqueda.

Los nodos de un directorio generalmente son más pequeños que los del archivo central y por esta razón, es más fácil rastrearlos; es necesario tener un nodo por cada elemento que forma el Banco de Datos, desde el mismo momento en que se crea. Si el nodo se aloja en el directorio y no en el archivo central, el espacio de memoria que se desperdicia es sensiblemente menor.

Cuando un archivo es de longitud variable, la única forma de accederlo es a través de los directorios en los que generalmente se tienen registros lógicos que contienen un apuntador a la posición donde se inicia el nodo de información en el archivo y algunos datos complementarios que permiten seleccionarlo.

Estos directorios permiten el acceso a la información a partir de llaves numéricas o a partir de una llave formada con -

la configuración exacta de bits de una palabra.

Directorio numérico del archivo central.

Se utiliza para guardar la información de cada ficha que permite localizarla y manejarla fácilmente. Por cada ficha del archivo central, existe un nodo de tamaño fijo en el directorio numérico que se localiza a partir del número de tarjeta y se compone de los siguientes campos.

a). Directorio de la ficha en el archivo central.

La dirección al archivo central indica exactamente el registro y la palabra donde se encuentra el inicio de la ficha.

b). Información particular de la ficha.

Aquí se localizan los datos que son útiles para seguridad del sistema y son:

- Número de archivo de captura en donde se encuentra la ficha (es un número secuencial que se asigna a los archivos de entrada cada vez que éstos se dan de alta en el Banco de Datos), sirve para localizar el documento fuente en caso de requerir una -

aclaración al contenido del mismo.

- Clave del origen de captura.- Indica si la ficha se capturó en la Dirección General de Bibliotecas o por proveedores.

c). Dependencias que cuentan con el libro.

Esta parte del nodo se compone de 240 bits, cada uno re presenta una de las bibliotecas que pueden tener el título en su acervo.

El archivo tiene registros físicos de 240 palabras; el registro cero se usa para los controles y los nodos de información se alojan a partir del registro uno.

Los arreglos de trabajo utilizados tienen seis palabras más - que el tamaño del registro físico, las cuales se utilizan para el control del mismo. La información que contienen estas palabras se encuentra en el registro cero y se transcriben en el momento de abrir el archivo.

Cuando el archivo se actualiza, la información de control se regrababa en el registro cero al momento de cerrarlo.

La información contenida en las seis palabras de control del archivo es:

a). Número de registro.

Indica que número del registro del archivo se encuentra cargado en el arreglo.

b). Actualización.

Un valor diferente de cero indica que el registro cargado en el arreglo fue actualizado y requiere copiarse al archivo.

c). Límite superior.

Indica cual es el valor más grande del número del dato que puede utilizarse en el directorio.

d). Tamaño de los datos.

Indica cuál es la longitud en caracteres de cada dato.

e). Número de datos por registro.

Señala cuantos datos caben en cada registro.

f). Contador secuencial.

Indica cual es el próximo dato que entra en secuencia.

g). Caracter.

Es el caracter con que se llena cada nuevo registro al expandirse el directorio.

h). Tipo de archivo.

Se utiliza para corroborar de que tipo de archivo se trata.

i). Abierto.

Indica que el directorio fue abierto y por lo tanto se puede utilizar.

Directorio de llaves.

Después de alojar físicamente los datos en el archivo central se procesa la información para actualizar los directorios de llaves.

Se relacionan las cadenas de cada una de las fichas que se utilizan como llaves para localizarlas directamente y se escogen los campos que se consideran relevantes para la recuperación.

Las llaves se forman de dos maneras:

a). Estricta.

Se utiliza toda una frase respetando el orden de los -  
elementos que la forman.

b). Libre.

Se escogen las palabras significativas de los campos,  
eliminando los artículos, preposiciones y conjunciones;  
se normalizan y se les aplica una transformación foné-  
tica para evitar los probables errores ortográficos.

Los elementos seleccionados para formar llaves de recupera-  
ción son:

- Número de clasificación.
- ISBN (International Standard Book Number).
- Autores.
- Títulos.
- Temas.
- Editorial.
- Serie.

Después de haber normalizado la información, se le aplica una función de dispersión (HASH) que la transforma en un número. Esta función ha sido seleccionada después de muchas pruebas - de concatenación de bits, se usa para dos fines:

- 1.- Para formar la llave de identificación en el Directorio de llaves únicas.
- 2.- Para localizar el nodo en el directorio de llaves únicas.

Existen dos tipos de directorios de llaves, a saber:

- 1.- Directorio de llaves únicas.
- 2.- Directorio de listas de llaves.

#### Directorio de llaves únicas.

Cuando se aplica una transformación a una palabra, convirtiéndola en llave y es nueva, se almacena en este archivo asociando esta llave a la del archivo de directorio de lista de llaves.

Como llave, se usa la dispersión de las listas recuperables y como número asociado, el apuntador al archivo de directorio - de listas.

Cuando se trate de almacenar una nueva llave en un registro - de trabajo y éste esté lleno, se utilizan registros de sobre- flujo para alojarla junto con su número asociado.

Una aplicación de este tipo de archivo es utilizarlo para con tener los apuntadores a listas de recuperación de fichas bi- bliográficas.

Tanto los registros de área normal de trabajo como del área - de sobreflujo se utilizan registros de 240 palabras.

En cada registro se alojan 159 llaves con su número asociado y se componen de tres secciones:

- 1.- Sección de llaves.
- 2.- Sección de números asociados.
- 3.- Sección de control.

La información contenida en el registro de control es:

a). Número de registros.

Indica qué número del registro del archivo se encuentra cargado en el arreglo.

b). Actualización.

Un valor diferente de 0 indica que el registro cargado en el arreglo fue actualizado y requiere copiarse al archivo.

c). Area normal.

Indica cuántos registros forman el área normal de trabajo.

d). Sobreflujo.

Indica cuantos registros forman el área de sobreflujo.

e). Llaves contenidas.

Indica cuantas llaves diferentes tiene el archivo.

f). Llaves de sobreflujo.

Indica cuantas llaves diferentes estan guardadas en el area de sobreflujo.

g). Tipo de archivo.

Se usa para corroborar de que tipo de archivo se trata.

h). Abierto.

Indica que el directorio de llaves únicas fue abierto y por lo tanto, se puede utilizar.

#### Directorio de listas de llaves.

Para cada llave que se obtiene, es necesario construir una lista, cuyos elementos contienen referencias y datos específicos de las fichas en que aparecen.

Estos permiten tener una rápida relación de todas las fichas que contienen una llave en común.

Para que la actualización del archivo de listas sea rápida, se ordenan las palabras que se van a dar de alta, de esta manera puede haber un bloque de palabras que se almacenen en lugares contiguos en un mismo registro.

#### Archivos auxiliares.

En una Base de Datos bibliográfica como el sistema LIBRUNAM,

el volumen de información es muy grande y no es adecuado almacenar información repetitiva, es decir; cuando se tienen varios volúmenes de un mismo título, únicamente se debe guardar un solo registro y el control de los otros volúmenes debe almacenarse en estos archivos auxiliares.

Además se tienen que almacenar estadísticas por períodos de los procesos e información que se utiliza por éstos.

Una de las características de las bases de datos es su seguridad en cuanto a la pérdida de información, para esto, se tienen archivos que guardan una imagen de la información que se da de alta en el Banco.

Los archivos auxiliares que se utilizan en el sistema LIBRUNAM, son:

1.- Archivos mensuales de adquisiciones.

La información de estos archivos se guarda en forma secuencial cuando se actualiza el Banco y al final del mes se protegen en una cinta magnética, los datos que contienen son:

- a) Etiqueta de tipo de proceso.

\$035 indica que es título original, \$036 indica -  
que es de cargo a una dependencia.

b). Número de ficha.

Es el número que identifica el registro bibliográ-  
fico en el banco de datos.

c). Etiqueta de dependencia.

\$900 indica el nombre de la dependencia.

d). Código de Dependencia

Es la clave de la dependencia que obtuvo el volumen.

e). Etiqueta de adquisición.

\$901 contiene el número de adquisición que le co--  
rresponde un determinado título.

f). Número de adquisición.

Es el que controla el acervo existente en cada de-  
pendencia, debe ser un número único para cada fi--  
cha.

Este archivo se utiliza también para generar estadísti-  
cas mensuales de producción, elaboración de catálogos,  
además de controlar y formar el inventario departamen--  
tal.

2.- Archivo de imagen de fichas bibliográficas.

El objetivo de este archivo es tener una copia de la información del Banco de Datos ya depurada para casos extremos de pérdua o un siniestro en la computadora.

Estos archivos se cierran cada mes protegiéndose en cintas magnéticas.

3.- Archivos de estadísticas.

Se generan archivos con datos de estadísticas y estadísticas ya procesadas mensualmente. La información contenida es la siguiente:

- a). Proceso original.
- b). Proceso de cargo.
- c). Idiomas.
- d). Clases de clasificación.
- e). Areas.
- f). Tipo de publicación.

g). Fechas de publicación.

h). Dependencias.

#### 4.- Archivo de dependencias.

En este archivo estan almacenadas las claves de las dependencias existentes en la Universidad Nacional Autónoma de México.

Cuando se da de alta uno o varios registros bibliográficos, existe una rutina que checa que la clave de dependencia esté correcta y entonces se podrá almacenar en la base de datos.

#### Archivos temporales.

En los procesos de actualización de la Base de Datos, así como de la explotación de la misma, se utilizan algunos archivos temporales que sirven para validar, clasificar, normalizar, etc. la información.

ESTRATEGIA PARA EL MANTENIMIENTO Y EXPLOTACION DE LA BASE DE  
DATOS LIBRUNAM.

Problemática.

Actualmente para generar la producción normal del sistema - LIBRUNAM, ha sido necesario efectuar correcciones a algunos - procedimientos básicos, lo cual se dificulta debido a los problemas siguientes:

- a). El sistema está diseñado de una forma muy compleja, lo cual dificulta alternativas posibles de futura expansión o cambios.
- b). Pruebas incompletas del sistema.
- c). Cambios en los requerimientos al sistema, debido a que no existe documentación de acuerdos en el diseño del - sistema y dado los cambios en la estructura de la organización de la Dirección General de Bibliotecas.
- d). Algunos procedimientos han sido programados sin un estudio profundo dado por las limitaciones de tiempo, personal y recursos de equipo, por lo que se dificulta la generación de producción normal y especial.

- e). Avances técnicos de Hardware y Software.
- f). No existe documentación específica de ciertos procedimientos.
- g). Rotación y adiestramiento de personal.

Soluciones propuestas.

Para optimizar el funcionamiento adecuado del sistema LIBRUNAM se proponen las siguientes estrategias a seguir, mismas que - facilitarán la explotación máxima de los servicios potenciales que estan previstos en el sistema.

- 1.- Auditar el sistema para asegurar que opere de acuerdo a las especificaciones originales y evaluar cambios propuestos. Para realizar dicha auditoria es necesario -- procesar cada una de las rutinas que contiene LIBRUNAM revisando los resultados que genere y comparándolos con los objetivos establecidos.
- 2.- Corregir ciertos errores de lógica del programa para - asegurar la buena interrelación de las especificaciones.

- 3.- Integrar ciertos procedimientos que estan operando de manera aislada. Es necesario hacer un análisis de las rutinas que se encuentran aisladas, de tal manera, que funcionen adecuadamente en la aplicación que realiza.
- 4.- Crear la infraestructura capaz de brindar la totalidad de los servicios en una forma organizada y consistente a las dependencias que integran el sistema bibliotecario universitario, así como a otros usuarios potenciales del sistema en un radio nacional e internacional.

Para crear la infraestructura adecuada, se utilizará la siguiente metodología:

- a). Analizar las necesidades y demandas de los usuarios a modo de integrar servicios destinados a subsanarlas.
- b). De acuerdo con los requerimientos de los usuarios y los recursos existentes, tanto humanos como técnicos, se diseñarán o modificarán rutinas en el sistema LIBRUNAM para satisfacer estos requerimientos.
- c). Asesoramiento tanto al personal que se encargará

de brindar dichos servicios así como a los usuarios de éstos.

- 5.- Generar la documentación complementaria del sistema, enfocada tanto al personal de cómputo como a los usuarios.

La documentación destinada al personal de cómputo, explicará en forma detallada la lógica general de los diferentes comandos de aplicación y de las rutinas específicas que componen estos comandos los cuales se han modificado y se ha creado algunos nuevos.

También se indicará de una manera sencilla y clara la forma en que funcionan los llamadores de las aplicaciones.

La documentación enfocada a los usuarios, detallará el uso de las distintas aplicaciones que brinda el sistema LIBRUNAM.

NECESIDADES DE SOFTWARE Y HARDWARE PRESENTES Y A UN PLAZO DE DIEZ AÑOS.

Software. Necesidades presentes.

- a). Corrección de algunas rutinas del sistema.
- b). Integración de rutinas aisladas al sistema central.
- c). Desarrollo de un procedimiento para la regeneración del archivo de dispersión y listas.
- d). Modificar las rutinas existentes de producción normal - para su funcionamiento automático a través de la programación de llamadores adecuados.
- e). Análisis y diseño de rutinas que contemplen la máxima - explotación de la Base de Datos, brindando la totalidad de los servicios previstos.
- f). Para la optimización de recursos, es necesario integrar rutinas básicas del sistema LIBRUNAM a la biblioteca de la computadora B6700, lo cual significa una reducción - en el tiempo de respuesta y el espacio en memoria.

- g). Análisis y diseño del nuevo módulo de catálogos de autoridad que permita uniformar la información contenida en el Banco de Datos.

Software. Necesidades a diez años.

- a) Actualmente a nivel nacional e internacional, se ha solicitado el software de LIBRUNAM pero como éste está programado en un lenguaje que no está en todas las computadoras, será necesario hacer una conversión del mismo a otro lenguaje más comercial; por lo tanto, hay que analizar que lenguaje sería el más apropiado para poder ser transferido a cualquier computadora y que tenga el software necesario para el proceso de conversión.
- b) En la base de datos se ha incurrido en errores de ortografía, errores técnicos y duplicidad, debido a lo cual se están desarrollando los catálogos de autoridad a partir de algunos elementos contenidos en la ficha bibliográfica y la fase final será hacer toda la estructura en software para regenerar todos los archivos de LIBRUNAM en base a los catálogos de autoridad.
- c) El Banco de Datos LIBRUNAM, se actualiza semanalmente y

el crecimiento cada vez está haciendo más lentas las investigaciones en línea, por esta razón, es necesario haceer un estudio para dividir los archivos ya sea en áreas de la ciencia o en elementos de la ficha bibliográfica; esto es, si normalmente para el servicio a los usuarios, la localización se hace a través de los temas y ya que en la actualidad se tiene todo junto en un mismo archivo, se desperdicia tiempo y espacio lo cual no sucedería si se tuvieran divididos los elementos de los archivos lo que permitiría acceder el Banco a través del perfil deseado.

Hardware. Necesidades presentes y a diez años.

La información bibliográfica que se da de alta en el Banco de Datos LIBRUNAM es de 8,000 fichas bibliográficas mensuales, - aproximadamente que forman un total de 16,000 segmentos.

A la fecha, los archivos que forman el banco de datos ocupan 800,000 segmentos aproximadamente y los archivos de trabajo - temporales, llegan a extenderse hasta 60,000 segmentos.

Por otra parte, la capacidad de las unidades de packs en la -

computadora son de un millón de segmentos, esto indica que únicamente hay capacidad de almacenamiento para seis u ocho meses aproximadamente.

Además, el número de usuarios que comparten los recursos de la computadora, es demasiado para la capacidad de la misma y dichos usuarios siguen incrementándose, lo que trae como consecuencia que el tiempo de respuesta sea cada vez más lento y la producción del sistema LIBRUNAM se retrase constantemente.

Está en proyecto, un proceso de desconcentración bibliotecario en el cual se pretende llegar a una catalogación cooperativa que conlleva un fuerte incremento en los procesos de apoyo automatizado.

En base a las necesidades de software y a las limitaciones de Hardware, es necesario, para poder brindar un apoyo eficiente, tanto al departamento de procesos técnicos como a los usuarios del Banco de Datos LIBRUNAM contar con los recursos de equipo y humanos que posteriormente se mencionarán.

Recursos de Equipo.

Base de datos.	2,000,000 Segmentos
Desarrollo de Programas.	80,000 Segmentos
Archivos de Pruebas.	40,000 Segmentos
Archivos Temporales.	100,000 Segmentos
Archivos de Tránsito.	10,000 Segmentos
Archivos para catálogos especiales.	1,000,000 Segmentos
Terminales.	100 Terminales

Recursos Humanos.

- 4 Analistas Programadores.
- 2 Investigadores.
- 1 Contador
- 2 Técnicos Académicos.
- 3 Codificadores.

B I B L I O G R A F I A .

Burroughs Corporation. Information Processing Systems Reference Manual. Burroughs Corporation. Detroit, Michigan: Burroughs Corporation, 1972. Pag.varía.

Lancaster Wilfrid F. The Measurement and Evaluation of Library Services. Washington, D.C.: Information Resources Press, 1977.

Librunam: Sistema Automatizado para Bibliotecas por Charlotte Bronsoiler... [et. al] UNAM, Dirección General de Bibliotecas. México: UNAM, D.G.B., 1979. pag.223.

