

S. TENZER, R. LAVEGA

RESUMEN

Se implementó un sistema computarizado de apoyo para el almacenamiento, la clasificación y la recuperación, total o selectiva, de información bibliográfica. Los procesos interactivos son en tiempo real y operados por bibliotecarias. La consistencia de información se lleva a cabo mediante un autómata finito determinístico. Las consultas interactivas se soportan a través de cuatro estructuras de datos indexadas eficientes.

ABSTRACT

A computerized support system for storing, sorting and retrieving bibliographic information has been implemented. The librarians perform by themselves the on-line interactive data entry and consultations. The Information checking is performed through a Finite Deterministic Automaton. The interactive retrieving is based on four efficient indexed data structures.

El Centro Latinoamericano de Perinatología y Desarrollo Humano - CLAP, desde hace quince años, trabaja en investigación, docencia y cooperación técnica con los países de la Región de las Américas, en el campo de la perinatología y el desarrollo humano. Para lograr sus objetivos, el Centro posee una biblioteca especializada, en la cual el manejo manual de la información bibliográfica era el problema principal. En particular, no había índices disponibles sobre las propias publicaciones del CLAP, que superan el millar.

Debido al grado de especialización de la Biblioteca los índices internacionales, tales como por ejemplo los confeccionados por BIREME - Biblioteca Regional de Medicina y Ciencias de la Salud (San Pablo) no satisfacen las necesidades del Centro.

Desde 1980, el CLAP tiene su propia computadora, una PDP11/34, para todos sus propósitos: manejo de los bancos de datos perinatales, procesamiento de información epidemiológica, investigación, docencia de estadística e informática aplicada a la salud, etc. Se decidió hacer uso de la computadora para el tratamiento de datos no-numéricos a efectos de contar con un sistema para clasificar y recuperar la información bibliográfica.

IMPLEMENTACION

El Sistema se implementó a lo largo de varios años, agregándose sucesivamente nuevas facilidades.

Para definir la implementación fue necesario analizar, estudiar y documentar todo el trabajo de la Biblioteca, con énfasis en la selección de entradas bibliográficas y la indexación de los documentos, que incluyen: publicaciones periódicas, apartados, informes, tesis, libros, congresos y repartidos. Sólo en apartados hay más de 20,000.

Se diseñó el sistema y se estableció que la transcripción a medio magnético de las entradas bibliográficas fuera mediante reglas muy similares a las empleadas en bibliotecología.

Las entradas bibliográficas se llaman asientos. Cada asiento contiene los datos requeridos para identificar y clasificar un documento. Ellos son:

- Código de publicación
- Autor y coautores
- Título
- Publicación (pie de imprenta y colación)
- Notas:

- Palabras claves (descriptores)
- Referencias
- Resumen
- Otros
- Forma en que se encuentra en la Biblioteca

El código de publicación identifica al tipo de documento clasificado, el cual puede ser: libro, parte de un libro, congreso, parte de un congreso, tesis, artículo de publicación periódica, informe o repartido.

Los programas fueron codificados en lenguaje COBOL, específicamente COBOL Plus (con CLASS para el diseño de las pantallas) bajo sistema operativo TSX. Actualmente el Sistema se compone de 79 programas.

Las pantallas para tareas interactivas se han diseñado en forma sencilla y cuidando la estética. Además, el pedido de campos asegura la especificación válida de parámetros.

El Sistema cuenta con módulos interactivos a cargo de las propias bibliotecarias y otros en tiempo diferido, a cargo de operador de computadora.

Los módulos en línea están guiados por menús que aparecen en la pantalla de la terminal de video, permitiendo que las bibliotecarias realicen su tarea de manera cómoda y segura.

MODULOS

Los módulos del Sistema son los siguientes:

Propósito:	Modalidad:	A cargo de:
Ingreso y corrección de asientos	ON LINE	Bibliotecaria
Consistencia de asientos y otros controles	BATCH	Operador
Emisión de índices impresos	BATCH	Operador
Definición del diccionario de descriptores	ON LINE	Bibliotecaria
Generación de estructuras para consulta interactiva	BATCH	Operador
Consulta interactiva	ON LINE	Bibliotecaria
Estadísticas	BATCH	Operador
Tareas complementarias	ON LINE	Bibliotecaria

En el módulo de ingreso y corrección de asientos, las Bibliotecarias, se encargan de la preparación de los datos. Para ello cuentan con un menú en pantalla, un programa editor de textos y el manual de reglas de transcripción.

El módulo de consistencia de asientos y otros controles es operado periódicamente. Se realizan los controles de los asientos ingresados y corre-

gidos, y se incrementa el banco de datos con los correctos. Los controles son:

- sintaxis de los asientos (aplicando técnicas de teoría de autómatas)
- asientos con número de inventario repetido
- detección de posibles asientos duplicados (por semejanza en texto)
- descriptores no definidos.

Los errores son corregidos por las propias bibliotecarias.

El módulo de emisión impresa de índices clasifica todos los asientos según el concepto de interés y presenta la información en un único formato tipo. Ver figura 1.

Se emiten tantas copias como se deseen o bien un original, que sirve para la impresión a mimeógrafo o en offset. Los índices disponibles son los ordenados por: Nro. de inventario, Título, Autor, o Decreciente por fecha y autor. También se especifica la inclusión, o no, de los resúmenes.

CLAP
INFORME DEL CLAP 1970-1975. INVESTIGACIONES REALIZADAS.
OPS, REUNION DEL COMITE EJECUTIVO DEL CONSEJO DIRECTIVO, WASHINGTON, JUNIO, 1976
37P.
APARTADO NRO. INV. 1067700 I - INFORME

CLAP
NARRATIVE REPORT OF THE ACTIVITIES ACCOMPLISHED BY THE
CLAP.
OPS, WASHINGTON, , 1973
APARTADO NRO. INV. 1057900 I - INFORME

CLAP
INFORME DE LAS ACTIVIDADES REALIZADAS EN EL PRIMER
SEMESTRE DE 1979.
OPS, WASHINGTON, 1 EN. - 30 JUNIO., 1979 33P.
APARTADO NRO. INV. 1083300 I - INFORME

COCH, J.
DEVELOPMENT OF A NEW METHOD FOR THE ISOLATION OF
OXYTOCIN FROM HUMAN BLOOD.

OPS
RESEARCH IN PROGRESS. 1974.
WASHINGTON, OPS, 1974, P.132-134

REF.
L'W20.5-PANR NRO. INV. 1059210 LP - LIBRO-PARTE

COLLIA, F.
DIAZ, A. G., FESCINA, R., HORCHER, R.,
RIMOLI, R., SCHWARCZ, R., TENZER, S.,
URMAN, J.
DIAGNOSTICO DE SITUACION PERINATAL. RESULTADOS
OBTENIDOS CON LA APLICACION DE LA HISTORIA CLINICA
PERINATAL EN EL HOSPITAL SARDA DE BUENOS AIRES.
MONTEVIDEO, CLAP, 1978, 39P.
APARTADO NRO. INV. 1078600 L - LIBRO O FOLLETO

CORDANO, M. C.
CRECIMIENTO FETAL INTRAUTERINO EN EMBARAZOS GEMELARES.
PATRONES ECOGRAFICOS.

MONTEVIDEO, URUGUAY CLAP, 1981, 97P.

TESIS DE DOCTORADO, REF., RES.

APARTADO NRO. INV. 1093500 T - TESIS

EL CONOCIMIENTO DEL CRECIMIENTO FETAL ES UNO DE LOS CONCEPTOS BASICOS QUE DEBE
DOMINAR TODO PROFESIONAL QUE TENGA LA RESPONSABILIDAD DE VIGILAR UN EMBARAZO.

EL ESTUDIO DEL FETO COMIENZA CON EL ESTUDIO DE LOS PATRONES DE
CRECIMIENTO. ESTOS DEBEN SERVIR COMO GUIA PARA LA DETECCION, DIAGNOSTICO Y
TRATAMIENTO DE LOS TRASTORNOS DE LA VIDA INTRAUTERINA.

Y CUANDO NOS ENCONTRAMOS FRENTE A UN EMBARAZO DE ALTO RIESGO POR SU GRAN
CONTRIBUCION A LA MORTALIDAD PERINATAL Y POR LA FRECUENTE MORBILIDAD FETAL QUE
ENTRAÑA, COMO LO ES EL EMBARAZO GEMELAR, LO ANTERIORMENTE DICHO COBRA AUN MAYOR
IMPORTANCIA.

EL PRESENTE ES UN ANALISIS PROSPECTIVO, QUE CONSISTE EN UN ESTUDIO DE TIPO
LONGITUDINAL DE EMBARAZOS GEMELARES, REALIZADO EN LA UNIDAD DE INVESTIGACION Y
DIAGNOSTICO ECOGRAFICO DEL CLAP, DE 1979 A LA FECHA.

FUERON SEGUIDOS SEMANALMENTE DEL COMIENZO AL TERMINO DE LA GESTACION, 32
PACIENTES SANAS, PORTADORAS DE EMBARAZO GEMELAR, EN LAS QUE SE MIDIERON
PARAMETROS CRANEANOS, ABDOMINALES Y RELACIONES ENTRE LOS MISMOS.

SE ANALIZARON LOS DATOS OBTENIDOS DE TODA LA POBLACION Y DE DOS GRUPOS
SELECCIONADOS SEGUN PESO AL NACIMIENTO, GEMELARES MAYORES Y MENORES.

FUERON COMPARADOS ESTOS DOS ULTIMOS GRUPOS ENTRE SI Y CON LOS PATRONES DE
CRECIMIENTO OBTENIDOS DE EMBARAZOS UNICOS.

LOS RESULTADOS OBTENIDOS MOSTRARON UN CRECIMIENTO DE GEMELARES MAYORES, EN LA
MAYOR PARTE DE LOS INDICADORES, ASIMILABLE A LOS RESULTADOS DE LOS UNICOS.

LOS GEMELARES MENORES SE COMPORTARON COMO RETARDOS EN EL CRECIMIENTO
INTRAUTERINO.

A TRAVES DE LA MONITORIZACION INTRAUTERO DEL CRECIMIENTO FETAL MEDIANTE LA
TECNICA ULTRASONOGRAFICA, QUE PERMITE LA REALIZACION DE MEDIDAS DE FIELES
INDICADORES DE LA EVOLUCION FETAL, ES POSIBLE DETECTAR UN RETARDO EN EL
CRECIMIENTO INTRAUTERINO EN EL PARTO EN UNO DE LOS FETOS GEMELARES.

Figura 1: ejemplo de salida impresa.

El módulo de definición del diccionario de descriptores consiste en un subsistema para asegurar el uso de descriptores válidos. Actualmente la lista incluye casi 500 palabras claves. El propósito es mantener la mínima cantidad posible. El diccionario contiene sinónimos y comentarios. Ha sido confeccionado y se mantiene coordinado con BIREME.

Figura 2:

C . L . A . P .

26/03/85

C O N S U L T A B I B L I O G R A F I C A I N T E R A C T I V A

Opciones de Proceso.

- 1 - Ayuda (instrucciones operativas)
- 2 - Selecciona por Fecha, Autor y Descriptores
- 3 - Selecciona por Texto en el Título
- 4 - Visualiza trabajos seleccionados
- 5 - Clasifica índice de trabajos seleccionados
- 6 - Imprime los trabajos seleccionados
- 9 - Fin

Ingrese número de Tarea: _

El módulo de consulta interactiva permite la selección y la visualización en pantalla o en forma impresa de los asientos. Ver figura 2.

El usuario ingresa los parámetros de la consulta desde su terminal y el Sistema devuelve todos los asientos que son respuesta.

Las consultas son por:

- descriptores (hasta tres)
- autor y coautores
- fecha (en intervalo desde y hasta).

Pueden ser simples, especificando sólo un ítem de consulta o bien combinadas, especificando varios ítems, en cuyo caso la selección es tipo "AND".

Un Asiento puede estar clasificado bajo una, varias o ninguna palabra clave.

Puede contener un autor principal (no obligatorio) y uno, varios o ningún coautor. A los efectos de la consulta, el autor y los coautores se consideran indistintamente autores del asiento.

Las opciones de visualización de los asientos seleccionados incluyen (ver figura 3):

- impresión o despliegue
- asiento completo o sólo algunos campos
- recorrido ascendente por número de inventario o retroceso o bien al azar por número de inventario o por ordinal de selección.

Figura 3:

C . L . A . P .

26/03/85

CONSULTA BIBLIOGRAFICA INTERACTIVA

VISUALIZA trabajos seleccionados, SELECCIONADOS: 2

Intervalo de Fechas: 1970 A 1979 Autor: SCHWARCZ

Descriptor: RUPTURA PREMATURA DE MEMBRANAS
FETO

Número de Inventariò : 1059500 Ordinal: 00001

Autores:

SCHWARCZ	R.
BELIZAN	J, M.
NIETO	F.
TENZER	S.

P. claves:

AMNIOTOMIA
FETO
PARTO
RECIEN NACIDO
RUPTURA PREMATURA DE MEMBRANAS

Otros elementos:

LA ROTURA PRECOZ DE LAS MEMBRANAS OVULARES Y SUS EFECTOS SOBRE
EL PARTO Y EL NEONATO. INVESTIGACION COLABORATIVA ENTRE
MATERNIDADES DE AMERICA LATINA.

MONTEVIDEO,
CLAP,
1975,
80P.
RES,

A

(A-Ant. S-Sigte D-Directo)-Ver (R)-Reinicia (9)-Finaliza :

El módulo de estadísticas consiste en las salidas impresas que informan sobre la productividad en ingreso de asientos. También se generan estadísticas sobre la frecuencia de aparición de los nombres de autores y co-autores. En forma similar, se obtienen estadísticas sobre las palabras que componen los títulos de los documentos. Esta información es útil en la especificación de selección de asientos por texto contenido en el título.

El módulo de tareas complementarias se compone de subproductos obtenidos por las bibliotecarias, prácticamente sin intervención de técnicos en computación. Consiste en el uso de las facilidades existentes en el Sistema, como ser el editor de textos, para confeccionar cartas a los morosos, preparar bibliografías especiales, etc. Este es un resultado secundario importante a destacar por la iniciativa y autosuficiencia de las bibliotecarias.

AUTOMATA FINITO DETERMINISTICO

El control de consistencia de los asientos bibliográficos es realizado por una máquina de Turing combinada con un Autómata Finito Determinístico cuya fórmula es:

$$M = (I, O, Q, f, g, q_0, F)$$

M: máquina secuencial que acepta los asientos y genera una salida.

I: conjunto de los 64 caracteres de entrada (letras mayúsculas, números y caracteres especiales),

O: conjunto de 70 símbolos de salida,

Q: conjunto de posibles estados (son 80),

f: $Q \times I \rightarrow Q$ función de estado siguiente,

g: $Q \times I \rightarrow O$ función de salida en versión Mealy,

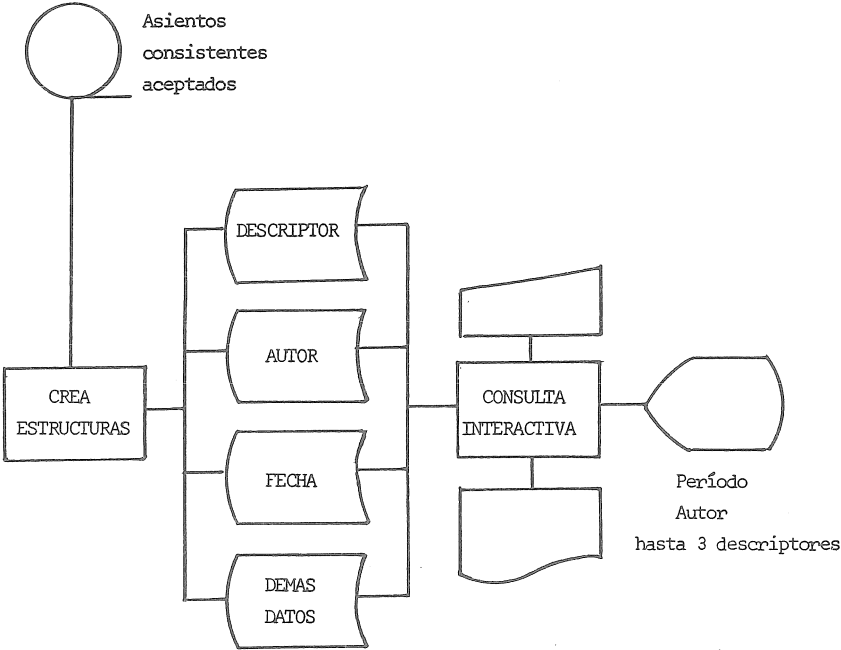
q_0 : estado inicial.

F: conjunto de posibles estados finales.

Las funciones "f" y "g" se almacenan en forma de matriz de transición, la cual puede ser alterada fácilmente, para incluir otros documentos o modificar las reglas de transcripción.

Debido a que la PDP11 con TSX permite time sharing con particiones de hasta 59 KB fue necesario dividir el programa en dos asincrónicos intercomunicados. Uno de ellos es la matriz y el otro el programa propiamente dicho.

Figura 4: Esquema de consulta interactiva



BASE DE DATOS PARA CONSULTA INTERACTIVA

Para permitir la consulta interactiva fue necesario crear estructuras de datos que soporten las relaciones de búsqueda y que sean eficientes en tiempo de respuesta y memoria.

A partir del archivo secuencial de asientos consistentes aceptados se generan cuatro estructuras de datos, indexadas (ver figura 4) que son:

- descriptores (palabras claves)
- autor
- fecha
- demás datos del asiento

El archivo de demás datos del asiento está compuesto sólo por aquellos elementos que no se encuentran en las estructuras anteriores. Se trata de introducir redundancia controlada, sólo la imprescindible, para ahorrar espacio en disco.

Cada asiento se identifica con un número de inventario único, que es el campo de vínculo entre todas las estructuras.

Las tres primeras estructuras, que son las de búsqueda, son lógicamente análogas y sólo difieren en contenido. Por lo tanto, a modo de ejemplo se sintetiza la correspondiente a descriptores.

ESTRUCTURA DE DESCRIPTORES

Un Asiento puede estar asociado a uno, varios o ningún descriptor. Un descriptor puede estar asociado a uno o varios asientos. La relación: Número de inventario - Descriptor no es biunívoca.

Se define así: Número de inventario -->>Descriptor

Descriptor ----->>Número de inventario

Permite accesos al azar y secuencial lógico por Descriptor y por Número de inventario.

Cada descriptor puede tener hasta 35 caracteres. Para reducir el factor de ocupación del disco se define como clave los primeros "n" caracteres, en este caso 15. El acceso es por un prefijo y por tanto se implementa una rutina de búsqueda con sinónimos.

En la estructura se mantienen cuentas de cantidad de asientos por descriptor. Entonces, para realizar la búsqueda el Sistema primero determina cual es la cadena (descriptor) más corta (que tiene menos asientos) para minimizar los accesos.

Ejemplo de consulta simple: el usuario selecciona por un sólo descriptor, todos los asientos clasificados bajo 'RECIEN NACIDO'. Se accede al azar por palabra clave y se recorre la cadena secuencial lógico. Todos los asientos de la cadena serán respuesta a la consulta.

Ejemplo de consulta combinada: el usuario selecciona por tres descriptores, todos los asientos clasificados bajo 'PESO AL NACER' y 'RECIEN NACI-

DO' y 'ALCOHOLISMO'. Se accede al azar por una Palabra Clave (la que tenga Cuenta menor) y se recorre esta cadena. Para cada elemento de la cadena se obtiene el número de inventario asociado, y se accede directo a la misma estructura por número de inventario investigando que las demás palabras claves de este Asiento cumplan con la condición. Si es así entonces el Asiento es respuesta a la consulta.

Con todos los asientos que son respuesta a la consulta se genera un archivo llamado Índice Respuesta. Contiene una entrada por cada Asiento que es respuesta, que básicamente es el número de inventario.

Se define así el vínculo para el acceso a todos los elementos que componen un asiento a los efectos de la posterior visualización.

CONCLUSION

El uso de la computadora en bibliotecología reduce el trabajo del personal, mejora la calidad de los datos, facilita la indización del material y su recuperación combinada, brindando servicios más adecuados a los usuarios.

BIBLIOGRAFIA

- 1) American Library Association. Reglas de Catalogación Angloamericanas. Washington, Organización de los Estados Americanos, 1970. (Manuales del Bibliotecario No.7)
- 2) Becker, J., Hayes, R.M. Information storage and retrieval: tools, elements theories. Somerset, New Jersey, Wiley, 1967.
- 3) Booth, T.L. Sequential machines and automata theory. Somerset, New Jersey, Wiley, 1967.
- 4) Jordan, J.R. Let the computer select your reading list. Datamation, 16(2):91-94, 1970.
- 5) Knuth, D.E. The art of computer programming. v.1. Addison Wesley, 1972.
- 6) Maegerlein, R.C. A simple-minded documentation system. Datamation, 16(2):131-132, 1970.
- 7) Matthews, F.W., Shillingford, A.D. Variations of KWIC. Aslib Proceedings, 25(4):140-152, 1973.
- 8) Salton, G. Automation information organization and retrieval. New York, Mc. Graw-Hill, 1968.
- 9) Starmer, F.C., Rosati, R.A., Simon, S.B. Interactive acquisition and analysis of discrete data. Computers and Biomedical Research, 5(5):505-514, 1972.
- 10) Tenzer, S.M. Análisis del sistema de información de la Biblioteca del Centro Latinoamericano de Perinatología y Desarrollo Humano. Montevideo, CLAP, 1973. (Publicación científica CLAP No. 574,01)
- 11) Tenzer, S.M. Automatización de la Biblioteca. Instructivo de transcripción de los asientos bibliográficos a tarjetas perforadas, Montevideo, CLAP, 1974. (Publicación científica CLAP 574,021)
- 12) Tenzer, S.M. Automatización de la Biblioteca, Reglas de transcripción de los asientos bibliográficos a computadora, Montevideo, CLAP, 1984. (Documento Interno CLAP 6/84)
- 13) Tenzer, S.M., Lavega, R., Uzal, M. Sistema de automatización de la Biblioteca. Sub-sistema ingreso de datos, Instrucciones de uso de la terminal de computadora, Montevideo, CLAP, 1983. (Documento interno CLAP 34/83)
- 14) Yerke, T.B., Computer support of the researcher's own documentation. Datamation, 16(2): 75-77, 1970.