DATABASE CREATION AND RETROSPECTIVE CONVERSION; A CASE STUDY OF INDIRA GANDHI MEMORIAL LIBRARY, UNIVERSITY OF HYDERABAD

E. Rama Reddy & R. Pommal Rao

ABSTRACT

This paper presents the Indira Gandhi Memorial Library's experience in Retrospective Conversion of its existing Catalogue records into machine readable form. It also describes how the database was created; updated and converted to LIBSYS format. Used structures for dbase database for the export of records to LIBSYS - Library application software and hurdles during the operation are discussed.

1. INTRODUCTION :

In the modern world timely, current, accurate and appropriate information is a prime commodity. The communication and processing of information has become one of the World's major endeavours. Scientific, industrial & technical areas are rapidly growing & simultaneously subjects are being expanded. New knowledge is acquired during the research & development programmes through out the world. Application of new technology has become essential. Computers have appeared on the scene in the history of scientific & technical development. Computers are becoming an indispensable tool in information processing and useful for multipurpose applications. Utilising Computer Technology is the only solution to the emerging needs of the clientele.

Library automation is to make library staff more productive with less effort. Library staff can provide the user needed material in time. Computer based library system is the best combination of good service & low cost.

If compared to human body, library is the heart of every organisation/institution/society; books are vessels & information is the blood. As blood circulation; purification is a continuous process & fresh blood is being added to the existing - similarly today's information will become retrospective for tomorrow. Every second, addition of current information is a must to keep up fresh the entire span.

Regular functioning can not be stopped, while additions are done. When the existing system is purified & added to regular stream it can be called **Retrospective Conversion**.

2. LIBRARY AUTOMATION :

The library is a catalyst/tool in the whole chain of social development. It is a dynamic store-house of information. All the major areas of library house-keeping functions can be mechanised. In a modern library there are so many tasks involved viz. Collection of Information; it's organisation; retrieval & dissemination, etc.

Quality of library services can be improved as computer methods become more widely adopted in libraries. All existing services should be updated or transfered to machine readable forms. One can not estimate the worth of automated library system, which is highly useful for future networking environment.

Four major areas recognised and well noted by the library professionals for computerisation in a library are:

> Acquisitions
> Cataloguing Process
> Circulation Transactions
> Periodicals Control

Retrospective Conversion is a part of Cataloguing Process, which covers the conversion of the existing data into machine readable form.

3. DATABASE :

The term "data" represents "information" in general. A database system is an integrated collection of related files, alongwith details of the interpretation of the data.

The automated Library itself is a big database consisting sub-databases like:-

> Books database
> Journals database
> Articles database
> Theses/Dissertations database, etc.

In this paper we have discussed about database creation, updation & retrospective conversion of earlier data related to books. Similar operations are used for the other databases. We have used the provisions of LIBSYS application software package through out the modules with our experience in building our Library database.

4.IGM LIBRARY EXPERIENCE :

After acquiring Intel 80286 PC/AT for the library automation in the year 1989, the first step was to design a database by using in-house developed dBase III plus application software. The structure of the database created is shown in **Annexure-I.** The titles processed for the entire month were entered into the main database LIBMAST. Printing a Monthly List of Additions along with catalogue cards and database searching for the required Author/Title/Key-Words **Accn. No.** Call No. was possible after indexing/sorting the database. Accession Index; Classification Index and Tracing Indexes were maintained for the purpose.

In the middle of 1991 LIBSYS package (Xenix based) was acquired; upgraded (Unix based) in 1993 on to Intel 80486 system and transferred to DEC Alpha 2000 client/server, DEC OSF based in mid 1995, now called Digital Unix, to suit the increasing user requirements. Simultaneously during 1991 the entire LIBMAST database (library catalogue) maintained through dBase III plus was exported to LIBSYS package.

4.1 DATABASE CREATION :

Before going in for creation of a database, it is necessary to decide on the following:

   # application of Cataloguing procedure
     (AACR-1/AACR-2/CCC)
   # application of Classification procedure
     (DDC/UDC/CC)

   # application of Computers
     a) Hardware required
     b) Operating System required
     c) Application Software to be used, and
     d) Staff training

The above factors depend on the collection of the library. Our library has 1,80,000 books; 58,000 back volumes and approximately 900 current periodicals. AACR-II and DDC 18th edition is being followed for the processing of books. Keeping in view the above facts, we have turned on to LIBSYS package which was mostly suitable for our library.

4.1.1. LIBSYS LIBRARY APPLICATION SOFTWARE :

LIBSYS package is written in "C" programming language and presented to the clientele with pre-designed modules.

> Acquisition
> Cataloguing
> Circulation
> Serials Control
> Article Indexing
> OPAC

Being a ready made & menu driven package, users cannot modify the field lengths (decrease or increase) and any user is not allowed to edit/customise the package. Vendor has been attending to customisation as per our requirements. In this way it is differing with CDS/ISIS package which can be controlled by a user for data base design & upgradation etc. We could discuss about the field lengths and sub-modules with LIBSYS vendor to suit to our needs. The fields in catalogue spread sheets are two types.

     a) fixed fields
     b) free fields.

4.1.2 DATA-ENTRY :

Keeping and maintaining error free database depends on Data Entry. Whether steps are taken to avoid incorrect entries/ etc. in the first level of the data entry itself will help the system administrator in multiple ways. Simultaneously we can spare the editing time for other purposes. In LIBSYS software .paediting is not very easy. Rotating the alert between the functions over the catalogue sub-module will become difficult.

> **Titles-in-Process :** After entering in to Cataloguing Module, first step is choosing **Titles-in-process.** Titles which are accessioned will appear in **Titles-in-Process** in Title alphabetical order. To go to

concerned title and to scroll the screens, use **F8** for continuation. Required title can be selected to fill-up the remaining bibliographical information and process it. Taking position to go to concerned title directly is not available at present. Entering Accession Number again is not necessary. If a particular title information is not found in the list, Enter-Title function can be selected.

Other way we can check through **Accession Number-Wise Display.** If the title is entered already by Acquisition Section, status will be shown as **in tech. processing** (in Acc. No. area of the Cat. card). For this type of titles going to **Titles-in-Process** will be compulsory to fill remaining details. If the title details are not entered in the data base, system indicates that the **specified Acc.No. is not Existing.** For this type of titles we choose **ENTER-TITLE** to fill the remaining columns/details.

Directly entering the data into **ENTER-TITLE** module without verifying the concerned title in **TITLES-IN-PROCESS,** will affect the data, which is entered through the Acquisition Section and it will lead to frequent updations, etc.

>Enter-Title : After opening this sub-module enter title & accession no. & press return key. System will check/verify the database for duplicate titles. If a title is shown as duplicate, check all the fields of the title to ensure that it is a duplicate. If the fields are matching we can add the accession nos. through **Update Holdings.** If not **F1** key is to be used to enter the data for new title.

4.1.3 DATA UPDATION :

The data entered through Enter-Title function above can be corrected by following the below mentioned steps. Titles which are found under issue status can not be used for data-correction. Changes will be allowed after discharging from circulation module.

> Update Title : is designed for corrections of spelling mistakes if any /to add extra information, which was forgotten at the time of cataloguing.

> Update Holdings : is used for adding 1. duplicate accession numbers **Volume No.** Copy No. & Location information 2. to remove or change existing Accession Numbers, etc.

> Remove Title Record : is useful to weed-out (remove) a title information completely from the database. Whenever a title consists multiple copies/volumes, it won't allow you to erase the record.

> Change Accession No. : here incorrect accession numbers can be corrected. Changing Prefixes (if any)

to Accession Numbers is also possible here.

4.2 RETROSPECTIVE CONVERSION :

> Definition : In general retrospective conversion means converting the back data into regular stream or database. Applying to catalog, we can say as "converting the catalogue data/record pertaining to previous titles, which were operated through Card Catalogue prior to the creation of database".

> Effect : Titles for which the catalogue cards printing and inclusion in the Monthly List of Additions is not required, may be entered through /Retrospective Conversion/. Titles entered through this sub-module will not appear in new arrivals file and card printing is also exempted. Only database will be updated and titles will be added.

> Operation : We can manage the Retrospective Conversion in two ways:

> a) entering records into Retrospective Conversion sub-module directly, and
> b) importing the records created in other than LIBSYS format - dBase / CDS-ISIS etc. (text format generated through dBase was followed here).

Our Library has been following Unit Card system since inception of the library. The shelf list cards are maintained in closed access and are being used for Stock Verification. We could complete stock verification of our stacks over a period of five years i.e. until July 1993. These shelf list cards are used for retrospective conversion since we have started our automation work from May 1989, all the new books purchased from May 1989 were already in the database. We could convert around 1,20,000 records under retrospective conversion ptoject. 12-14 library staff members worked in this project for over 7-8 months. They worked outside their working hours including holidays to complete this project under the supervision of an Asst. Librarian.

Since we have taken this work as a project with specific time frame, we have followed both the ways as mentioned above. Some of the P.Cs. were directly connected to the server & some were not connected. The connected P.Cs. were utilised for direct data entry into the LIBSYS format. Whichever was not connected were used for data entry through dBase operations..pa

We have used the structure as shown in **Annexure-I** for the creation of dBase database. After loading the records one by one into the dBase database, it was converted into a text file. Before taking a text file, the dbf file needs some corrections. Firstly fields

of the dbf file should be rearranged on par with the LIBSYS spread sheet design. The simultaneous order of the fields should be maintained. To overcome this problem, we have created another dBase file EXPLIB as shown in **Annexure-II.** Some areas are merged together and some are split in to 2 or 3. Records were transfered to EXPLIB. A text file was extracted from EXPLIB and copied into "wfilex" directory of LIBSYS. This text file is called input file for import. LIBSYS can not recognise the fields unless the above procedure is followed.

>To skip this procedure, in first level of dBase file itself we can create the structure on par with the LIBSYS.

Secondly, We have to note down the parameters used while loading dBase database, which are to be supplied at the time of importing records.

For example:-
Accession Number Field separators (;)
Accession Number separators   (/)
Key-Word separators   (-)
etc.

4.3 EXPORT & IMPORT :

In LIBSYS package under cataloguing module Import function is available. While loading the details into this function two options are to be selected and filled.

1. Import structure;
2. dBase structure (see **Annexure III**)

Supplied structures will be stored in two files with the name mentioned against "file name" with extensions **imp** & **txt** in **wfilex (working files directory)**.

>For importing of records, field lengths of the original dBase file has to be mentioned into no.2 above. In other option the details of fields, where those are available while tallying with this sheet, to be filled into no.1 above.

>Alongwith the above, input file type is to be filled. Some accepting file types are CDS/ISIS; Text; MARC and LIBSYS to LIBSYS Format. Instead of converting the records created through dBase III plus or dBase IV, CDS/ISIS file can be used.

>Import option can be selected as per our requirement. New records can be added/merged/replaced to the existing database.

>By using add function, titles will be added to the existing database. If any duplicate accession numbers are found in the dbf file, it will be kept separately by LIBSYS without accession nos.

>By using merge function titles will be merged in to the database. If the same titles are existing in the database the accession nos. loaded through dbf will be merged while import.

>While conducting the import, due to some system problems, we may have to run the repeat import operation for the same file. At that time using /replace/ option is compulsory. Replace function is little complicated. Some overwriting problems were detected after choosing **replace** for the test ride.

>System will display the status of the processed records alongwith duplicate records information and titles-without accession number details, etc at the end of the import session.

Following the above procedure step by step, we could import all the dBase files into the database of LIBSYS. For our convenience, we have taken an average 300 records at a time. If any problems detected after importing each file, it was easy to make corrections in that particular 300 records area. Checking for the incorrectly imported records is possible by executing **bibchk** available under **user** directory of LIBSYS.

4.3.1 HURDLES :

The process (a) discussed above in 4.2 is time-taking one. But one can load/make perfect entries.

In the second process (b) some hurdles are to be crossed.

>Converting original dbf file to EXPLIB
>EXPLIB file needs accession number
  corrections
>Field separators to be used correctly, for
  multi-volumesets, copies etc.
>File to be split in to pieces
>Problems during export/import
  - Unreadable characters
  - Problems with secondary storage
    devices
>After perfect import, copy statements &
  additional id.statements need updation
>In case of imperfect import, added
  records should be found and to be
  deleted for the re-running of the same file.

There were around 20% records found throughout the database, which needed data corrections later due to some of the hurdles at the time of retrospective conversion programme.
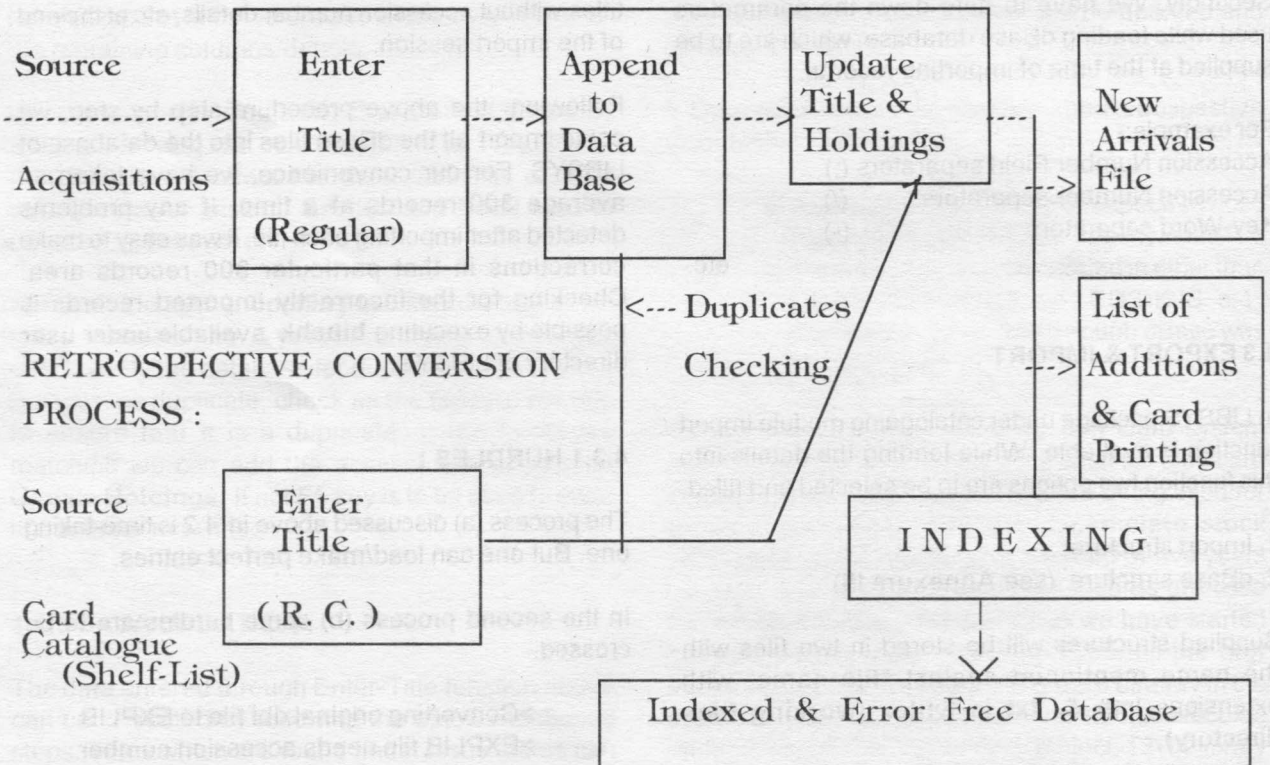
5 CONCLUSION :

Database creation including retrospective conversion is time consuming and need careful planning. Most libraries are worried about converting the older records bacause it requires money, manpower and proper decision making, etc. Libraries should creat their databases, which will enable them to get into the network and also provide better and efficient services to their users. UGC has recognised the importance of Library Automation and is providing special funding to libraries in phased manner to realise the objectives of INFLIBNET programme.

FLOW CHART:

CURRENT CATALOGUING PROCESS :

Source
Acquisitions

Enter Title (Regular)

Append to Data Base

Update Title & Holdings

New Arrivals File

<--- Duplicates Checking

List of Additions & Card Printing

RETROSPECTIVE CONVERSION PROCESS :

Source
Card Catalogue (Shelf-List)

Enter Title (R. C.)

I N D E X I N G

Indexed & Error Free Database

Structure for Database : LIB.dbf

Field	Field Name	Type	Width
1	GROUP	Numeric	7
2	COLL_TYPE	Character	1
3	CLASSNO	"	15
4	BOOKNO	"	6
5	AUTHOR	"	140
6	TITLE	"	180
7	EDITION	"	9
8	IMPRINT	"	63
9	V_P (COLLATION)	"	12
10	SERIES	"	100
11	NOTES	"	60
12	ACCN0	"	30
13	TRACING	"	150

Total width : 774

Structure for database : EXPLIB.dbf

Field	Field Name	Type	Width
1	TITLE	Character	140
2	SUB-TITLE	"	80
3	MAINAUTH	"	120
4	AUTH1	"	30
5	AUTH2	"	30
6	AUTH3	"	30
7	EDITION	"	9
8	PLACE	"	20
9	PUBLISHER	"	35
10	YEAR	"	4
11	V_P	"	12
12	SERIES	"	100
13	NOTES	"	60
14	ISBN	"	13
15	CLASSNO	"	15
16	BOOKNO	"	6
17	SUBJECT1	"	50
18	SUBJECT2	"	50
19	SUBJECT3	"	50
20	ACCNO	"	60

Total width : 914