Development of an information storage and retrieval system at a university library: a case study

Bashir Ahmad¹ and Nazir Ahmad Bhat²

Head, Library Services, Central Library, Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir, Shalimar, Srinagar, E-mail: bashirahmad555@gmail.com
Assistant Librarian, Central Library, Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir, Shalimar, Srinagar

Discusses the design and development of a database created for storage and retrieval of information about individual journal articles published in about 138 Indian and foreign current journals, 28455 journal issues existing in the back files section and full text of 356 theses available in electronic format at the Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir (SKUASTK). The database of books is managed through SOUL software and CD-ROM abstract databases running on WinSPIRS has been linked with this IS&RS. Entire IS&RS is being successfully run on LAN of the university. The IS&RS has been developed on MS Access platform with Visual Basic as front end. The IS&RS system consists of four modules viz: data entry, search, report, and administration-cum-monitoring. The database has been placed on the campus network and is running smoothly with encouraging operational results and interoperability with other databases.

Introduction

The Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir (SKUASTK) library comprises of central library at main campus Shalimar, Faculty of Agriculture Library at Wadura, Faculty of Veterinary Sciences and Animal Husbandry Library at Shuhama, Faculty of Fisheries at Rangel, Faculty of Forestry at Ganderbal, Regional Research Stations/Sub Stations located all over the Division of Kashmir including Ladakh. As of now university library system has around 90000 accessioned books, periodicals, theses and CD-ROM databases¹.

The university presently subscribes to 86 Indian and 52 foreign scientific journals. Till date around 28455 journal issues exist as back file collection. The university subscribes to 11 CD-ROM databases since 2003 with retrospective data from 1973, 1980, 1990, 1997 etc., which include CAB International, AGRICOLA, AGRIS, Biotechnology Abstracts, FSTA, ISA, MEDLINE, IPR, CARIS, VET CD and Bibliography on Rice. It has got a good supporting software WinSPIRS developed by Silver Platter/ Ovid Technologies, United Kingdom². Since 2003, the university has made it mandatory for the students to provide a soft copy of their theses and full text content. About 274 theses submitted by MSc. students and 82 by PhD scholars are in stock. With varied information sources as discussed, an attempt was made to develop an information storage and retrieval system for enabling information search and retrieval from a common interface.

Literature review

Before designing the IS&RS, a literature survey was conducted to ascertain information about the storage and retrieval systems presently in vogue in libraries of the country. It was found that majority of Indian academic and research libraries (i.e. 1688) are using the library automation software SOUL³, while around 1000 libraries use LibSys⁴ and the rest use others or no software at all. The trend to use DSpace, GreenStone, ePrints, Fedora Core software for development of digital libraries as full-text storage and retrieval systems has started recently⁵-⁶. It is evident from the literature surveyed that most of the Indian academic and research libraries are using the library softwares capable of storing and retrieving the information up to the bibliographical level only⁷-⁸. The libraries are not in a position to procure IS&RS to store and retrieve full length journal articles, books or that of theses etc, as their cost stands in the way⁹-¹¹. In this scenario, the libraries will have to use softwares that have capability of storing and retrieving information at bibliographical level and at full-text level. It may also be required to develop in-house integrated system to facilitate
seamless and instant access, use and management of information resources.

**Need for digitisation and development of a new IS&RS at SKUASTK**

Although library software namely SOUL (Software for University Library) was already in use, it suffered from some handicaps. SOUL has a provision for storage of bibliographical details of journals as a single entity. But this was not sufficient as it was required to index at the article level rather than at the journal level. Since the SOUL does not allow any customization at user level, it was not possible to integrate the WinSPIRS (software on which the CD-ROM databases run) with the SOUL. Also, it is not possible to place full text theses in the memo field of the SOUL that is primarily meant to accommodate the abstract of the thesis only.

Given the problems of having to use different software and interfaces to access different resources, a need was felt to develop an information storage and retrieval system that would be common interface for accessing the various e-resources of the library. Further, developing software on a platform/ application software on which the library staff has got some command leaves a room for future customization to meet the emerging requirements. As such the software developed indigenously eliminates dependence on outside agency for bringing about required day to day upgradation and setting right the system snags.

**Scope of digital contents**

The following items were included for access through the ISR&S.

i) **Journal articles**: Since the back files collection is now a considerable source of information, a need was felt to create a database of individual research articles contained in each issue of the existing journals. However, in order to give the project a practical shape, its scope was restricted by taking the journal issues from January 2005 onwards.

ii) **CD-ROM Databases**: 11 CD-ROM databases subscribed since 2003.

iii) **Theses**: 274 Theses submitted by MSc. Students and 82 by PhD Scholars, i.e. a total number of 356 theses comprising of around 46280 pages upto December 2007 and further day to day additions.

iv) **Link to SOUL**: A link has been created to have access to the bibliographical details of books and theses available on SOUL database. Here the users have been provided with a facility to browse the OPAC of the database.

**Design and development of database**

MS Access, relational database management system was chosen for developing the ISR&S.

**Under lying design**

MS Access consists of database objects for storage and retrieval of data. The data can be stored in tables and retrieved with the help of queries, filters, macros and modules. Forms are used to build different interfaces to encompass data entry and user-friendly retrieval interfaces. Moreover, one can design pages for subsequent posting on the website and linking the database with Internet. Six elements were identified with regard to bibliographical details stored which included:

1. Document/ source of information
2. Source of publication
3. Author(s)/ editor(s)/ compiler(s) etc.
4. Publishers
5. Data entry operators
6. Users

Accordingly six separate tables were used to store the data, namely Master, Sources, Authors, Publishers, Data entry operators and Users. Each table comprised of more than one fields. Every field was used to store information about constituent attributes of each entity. E.g. Table Master comprises of 6 fields i.e., MasterID (a unique key used to individualise and index the Table); Content (e.g. Title of the Article or Chapter); Text (e.g. Abstract or Full Text related to the content); Call No. and Page Nos. Likewise other tables comprise of more than one field to accommodate their respective attributes. Sub-fields were used to accommodate the sub-attributes associated with the Attributes e.g. (Author1; Author2; Author3 sub-fields of the main field Author(s)). The individual Tables were related to each other by creating the desired relationships. Three types of relationships were used where ever deemed fit to serve our purpose. Three types of relationships used where 1:1 (One to one:
only include rows where the joined fields from both tables are fully matching e.g. Journal Table : Publisher Table; One journal can have only one publisher; 1:m (One to many: Include all records from Child Table and only those records from the Master where the joined fields are fully matching e.g. Journal Table : Contents Table; One journal has got many article titles); m:1 (Many to one: Include all records from Master Table and only those records from the Child where the joined fields are fully matching e.g. Contents Table : Author(s) Table; Many articles might have been written by same Author(s)). The relationships were made possible using the concept of Primary Key and Foreign Key concepts. The creation of relationships is a pre-requisite to enable various Queries and Advanced Filters to function.

Navigation: As stated earlier forms have been used as interfaces to ensure logical navigation and dataentry. In this connection three Switchboards have been developed, namely i). Main Switchboard, ii). Search Plate and iii). Report Sheet.

**Main Interface**

The Main Interface (Fig. 1) possesses navigation buttons for a). Data entry (Dataentry of Journal Articles); b). Report Generation (View/Print Reports); c). Search (Search whatsoever); d). Performance Monitoring (Work Out Put of Dataentry Operators) and e). Stop Application (Exit The Database).

**Data Entry**

On double clicking the data entry button on Main Interface an interface developed on a form based on multiple underlying Tables gets activated (Fig. 2).

It enables the data entry operators to feed the relevant data into relevant fields of relevant tables behind the scene. So the data entry operators do not need to wander from table to table to make the data entry. Moreover, List Boxes and Combo Boxes have been used to pick and choose the standard data from the underlying tables containing such data kept in advance ready for future use. This is done to ensure uniformity in rendering the patterns of names, subject terms, keywords; thus avoiding any confusion in spelling and the corresponding syntax. Besides it, the Combo Boxes and List Boxes, save the typing time to a considerable extent. MS Access has got an advantage of using some functional keys and commands to maximize the dataentry speed. E.g. ^V

![Main Interface Screenshot](image-url)
Fig. 2 — Data entry sheet displaying the fields of the database for data input

(Control Ditto) - copies the data from upper immediate cell of the same field into the current cell in use. :
(Control Colon) - automatically puts in the current date and time in the cell in use and so on.

Search whatsoever

Double clicking this button on Main Interface activates the Search Plate (Fig. 3), the second switchboard possessing command buttons to meet various approach points of users for searching or locating specific pieces of information.

The main approaches through which the users try to search data are 1. Natural Language/Free Text; 2. Keywords/Keyterms; 3. Subject Headings; 4. Author Approach and 5. Combined Approach (complex query). Possibility for all the five approaches have been created through the five command buttons as shown in figure above. On clicking the Natural Language command button, a prompt (Fig. 4) asks for the word or word group one wants to search the information about. Once the word or word group is put in and OK is clicked the search engine browses the database and a list of the contents in the shape of hypertext gets displayed on the screen (Fig. 5).

The IS&RS has got the facility to get comprehensive bibliographical details of a particular content from the links (Fig. 6).

Report Generation

The software has got flexibility to generate tailored reports as per requirement (Fig. 7). At present a provision has been kept for generation of following reports: i) Content-wise Report; ii) Title of Document Wise Report; iii) Theme Wise Report. Fig. 7 shows the content-wise report.

WinSPIRS

Double clicking this button on Main Interface activates WinSPIRS, exclusive software to browse across the 11 CD-ROM Databases hosted by a separate Mirror Server.

Interoperability and standardisation

The software at present is operable on Windows operating system. The endeavour is on to make possible its interoperability with other reputed operating systems like Linux, UNIX, Macintosh, and the like. However, the data migration to other DBMSs is possible quite conveniently as the data is stored in Extended ASCII (8 bit) standard and can easily be transferred to Unicode or ISO 2709
Fig. 3 — Interface displaying search buttons to meet various approaches of the users

Fig. 4 — Use the search button to meet free text search on Title of the research article
Fig. 5 — Result of the example shown above to meet free text search on Title.

Fig. 6: Full bibliographical detail of individual article got as a result of the free text search on Title of the research article.

Fig. 7 — Report generation facility of the software.
files. The data available in MS Access objects i.e. Tables and Queries can easily be transferred to SQL or mySQL database objects. The later option will allow utilizing the stored data in open source Linux, Apache, mySQL, PHP (LAMP) atmosphere as well.

Conclusion
The commonly available database management system, MS Access has been effectively used to develop an information storage and retrieval system and also for creating a common interface to access the different electronic resources in the library. The database has successfully passed its initial test phase at SKUAST-K, main campus.

References
1. http://www.skuastkashmir.ac.in
2. Ibid
3. http://www.inflibnet.ac.in/service/soulsupport.html
10. Op cit Prem Chand and Chauhan
11. Op cit Cholin