By

B. T. Gangu

*Assistant Librarian*

*A-19, Professors Qrts*

*University of Hyderabad Campus*

*Gachibwli*

*Hyderabad – 500 046*

Yakub Ali

*Jr. Professional Assistant*

*A-19, Professors Qrts*

*University of Hyderabad Campus*

*Gachibwli*

*Hyderabad – 500 046.*

**Email: btglib@uohyd.ernet.in**

**Email: csbss@uohyd.ernet.in**

# ABSTRACT

The paper enumerates the innumerable technical problems encountered by libraries in managing the Urdu collection and the various hurdles encountered while switching over to an automated system from a manual environment. The authors try to impress the library professionals that it is high time they paid more attention to Indian Languages collections and also suggest how to over come the constraints through the examples of the IGM Library's experience.

**KEYWORDS: Urdu language collection, Cataloguing of Urdu collection, Database creation for Urdu language collection, University of Hyderabad Library**

## 0.     INTRODUCTION

The impact of information technology has been all pervading.  It has transformed the way the information is stored and disseminated and the traditional approaches to the library and its services are no longer adequate to meet the enlightened users of the present.  Thus the adoption of information technology is a necessity rather than a matter of choice for the modern libraries.  The National Educational Policy of 1966 as well as 1986 lay special emphasis on the study, teaching and research on all the Indian languages and its literatures along with English and Hindi. Therefore, the need for efficient organization of holdings and proper dissemination of information to the concerned user community. India has a rich and valuable collection of Urdu literature since Urdu was patronized by the Muslim rulers.  Much of the valuable manuscripts, as well as books languish in unsuitable, neglected atmosphere and could be lost to posterity.  Even the libraries in general have not been able to properly process, organize its Urdu collection due to lack of funds and technical tools to handle them.

## 1.	URDU LANGUAGE COLLECTION – TECHNICAL PROCESSING

The technical processing of Urdu language books is a bit complicated and requires a lot of care to properly process and organize the collection.  The widely used Dewey Decimal Classification has dumped Indian Languages and its literature under 490 "Other Languages", 890 "Literatures of Other Languages" respectively (and have to share this small portion with the languages of African and Asian origins).  Further, among the Indian Languages themselves, the multiple languages belonging to the Indo-Aryan and Dravidian groups have been squeezed into a small slot.  The Dewey Decimal Classification number for Urdu language is 491.439 (490: East-Indo-European and Celtic language; 491: Indo-Iranian (Aryan) languages.)

Thus, a class number building process for a Urdu language book (fiction or dictionary) involves a lot of time and concentration on the part of the classification expert to analyse and assign an appropriate number. Further, since the numbers tend to be too long, writing the number on the spine of the book, as well as shelving are problematic.  Brevity is the first casualty of the existing system of classification..

Since, standardized subject headings/ key words system are not available for Urdu language, the libraries have to generate their own key words or adapt the ones available for English language, such as Library of Congress, Sears List of Subject Headings etc.  In the case of automated system, the keywords have to be entered in roman script.  So the precision during the retrieval process suffers.

The cataloguing requires the adoption of the bibliographic entries in Urdu script to a standard code system. Preparation of cards, Main entry, added entries and the shelf list is also a laborious process, as they are to be manually prepared, unless of course a separate typewriter for Urdu language is available.

Since, the classe numbers are normally long, shelving of the language books cause confusion.  In the case of general books i.e. the Social Sciences and sciences etc., shelving becomes even more complicated as the language of the book is other than English language.

## 2.  MUSLIM NAMING PATTERN AND RENDERING OF NAMES

Mohammed Haroon in his book Cataloguing of Indian Muslim Names discusses in detail about the various problems connected with Muslim names and cataloging. In his opinion not much work has been done in the area of cataloguing of Muslim Names.  The first ever definition and detailed study on this subject came from Nasser Sharif whose book *Cataloguing of Persian Works* was brought out by American Library Association in 1959.

The only basis in this kind of treatment is the common denominator of the religious factor of Islam.  This basis of treatment is not sound.  The cultural traditions of countries which are now Islamic vary greatly.

The names of Indian Muslim authors can be found listed in Indian National Bibliography  (INB), Indian Books in Print (IBIP) and Library of Congress Accession list and these can be used as a source of information and a guide to cataloguers.  INB does not enter the Muslim names under a surname like Cheda or under a Nisbat as Rizvi, Ansari etc.  But often these are used as entry elements.  INB as a cataloguing tool is also not so efficient because libraries usually follow the use of surname and Nisbats as entry elements. IBIP is also not that useful as it does not provide the compiler a set of principles that can guide him.  Library of congress Accession List takes the last part of the Muslim names as the entry word and results in distortion of names.

### 2.1  Indian Muslim Naming Pattern and Rendering of Author's Name

Indian Muslim names generally consist of the following elements and in the given specified order.

·      KHITAB: It is an honorific title given to a person.  It is a dispensable  attachment to a name.  In early Muslim names, the Khitabs are found in abundance.  Khitabs are also given to many common names.

In India, the Muslim rulers conferred titles on some of their important subjects such as Military Generals, Warriors and Scholars.

*Cataloguing suggestions:*

      a) Khitab is a removable attachment

      b) Make entry under distinctive name

**c) If necessary, refer from Khitab.**

· **LAQAB:** It is a name of honour and is generally given by the public as Hakim, Mirza, Mulla, Moulvi etc. or the author may assumeit by himself. The name may begin with Laqab like words and yet be a part of the name as Hakim Ahemed, Amirali, Mir Alam, Sardar ali etc.

*Cataloguing suggestions:*

- It is removable attachment
- To avoid homonym in names, laqab may be used as an individualizing element:

o     Abdul Haq (Moulvi)

o     Mohammed Abdullah (Sheikh)

· **TAKHALLUS:** This is a literary name assumed by literary person in literature , and they are better known by this. It is of Persian origin and has come to India from there. Ranganathan considered the Takhallus as a literary name assumed by most men of letters and therefore treated it as a pseudonym. Thus Takhallus rendering is governed by the Rule MD42 and MD42 in Ranganathan's Classified Catalogue Code.

But Takhallus is quite different from pseudonym. In Urdu literature, it consists of one or two words and it usually indicates relationship to one's place of birth or residence, and is written at the end of a name. An author usually uses an pseudonym to conceal his identity and writes under pseudo name. But it is not the case with Takhallus, which originated from a custom of introducing at the close of certain poems, the poet's own name.

The reason for the origin of Takhallus by literary men and women was because they were interested in names that were popular and quickly identifiable. The same words will not be used as Takhallus by other contemporary writers. The Takhallus of Ghalib was Asad

Takhallus is the most easy way to remember a writer so that the library users approach the catalogue by Takhallus only.

*Cataloguing suggestions:*

**a) Takhallus should be the first choice for entry word.**

b) When full name of an author is known, it may be used for individualizing

and reference from it may also be given .

c). If Takhallus and Nisbat are present can be used as entry words.


INSHA (Saiyyed Inshaullah Khan)

JIGAR MORADABADI  (Ali Sikander)

SHAMIM KIRHANI (Shamsuddin Haider)


· NISBAT: In India, the Muslim community consists of Sunnis, Shias, Wahabis, Bohras, Qadianis, etc.. These are not castes but sects which follow different schools Islamic laws and theology.  Presently, there are about seventy two sects in India.


As far as the names are concerned Shaikhs, Sayyeds (Arabic), Beg , Khan (Turks) and Malik (Persian) were originally respectful honorifics.  Later on these came to denote the racial groups of the countries concerned. Among the converts, some retained their original surnames while the rest  adopted Nisbat names, denoting the relationship  of the person to a place, family, institution, sect, saint, tribe, etc.  Thus Nisbat is generally the last part of the name ending in 'i' or 'y'.


If there is a Takhallus in a given name, the cataloguer's search should be for identification of Nisbat or surname or a word performing the function of surname.  Dr.S.R.Ranganathan and INB do not favour Nisbat or surname as entry element.


*Cataloguing suggestions:*


For the purpose of cataloguing, Nisbat names may be favoured entry word as:


a)     The muslim authors after 1900 are becoming surname consciousness and more are using this.

b)     The author gives his Nisbat or surname on the title page and wants to be considered for cataloguing.

c)     Many surnames as Chaudhury, Patel, Sarkar, Sait are common to Muslims and non-Muslims. Cataloguers as well as readers are habituated to treat these as  surname.


·        Authors having surnames and family names should be entered under these

·        Authors having Nisbat names relating to caste, sect, family, saint and tribe should be entered under these entry element.

·        Authors having Nisbat relating to foreign places may also be entered under those place names.

·        Nisbat relating to places in India may be used for individualizing.

·        In case where more than one Nisbat qualify for entry word, all of them should be the entry words.

·        When an author uses a part of his name as surname, that part will be taken as entry word.


KOYA (C.H.Mohammed)

BANATHWALA (Ghulam mahamood)

NIZAMI (Khaliq Ahmed)

QUADARI CHISTI (Aziz Ahmed)


·        ISM: It is the personal name given to an individual at the time of his birth and consists of one or more characters for the following:


a) Ninety nine names of Prophet Mohammed

b) Ninety nine names of God

c) Names of other prophets mentioned in the Holy Quran

d) Names of other things etc. derieved from Arabic, Persian and

    local languages as Naseem, Jameel, Humayun, Iqbal, Nawab etc.


Muslim names consistings of Ism begin either with Mohammed and/or terminate as Ahmed, Ali, Hasan, and all such persons are better known by the other  parts of the name.  Ranganathan's canon of prepotents recommends the use of the most potent word in such names as entry word.


Example Mohammed Saeed, Saeed Ahmed , Saeed Ali where the word Saeed is the poten part and he is known by this name.  Whereas some names like Ahmed Mian, Ali Mohammed, Mohammed ali, Mohammed din, Din ahmed, din Mohammed are both words are potent and the authors are known by full name.  The other types of names like Zafar Iqbal, zafar Saeed, Zafar Mohammed are potent names only.


*Cataloguing suggestions:*

a) Two worded names consists of Ism should not be inverted.

b) If name has more than two words, then it will be

FAKHRUDDIN ALI AHMED

GHULAM HUSSAIN ZULFIQAR

c) If name has more words and begin with Mohammed, then it should be:

ABUL FAZAL FAROOQI (Mohammed)

ASLAM SAEED (Mohammed)

· **Compound Names**

 Many Muslim names begin with Abd or Abdul, God's name joined as Abdullah, Abdul. Such names written differently Abdul-ul-Rehaman, Abdul Rehamen, Abdulrehaman. This should not be inverted as it conveys some specific meaning.  Wherever it is to be used, full name should be joined by the element as al, an, as,e, i, ud, ul, ur, un, us etc. so that they can be easily recognized.

**2.2      Names of Muslim Women Authors**

Muslim women are known by their Kunyats or the name of relationship.  The name has Bi, Begum, Khatoon, Khanam, Bano and Nisa etc.as the second part of the name. Further  their father's name is replaced by the husband's name after marriage.

*Cataloguing suggestions:*

a) Simple women names as Nadira Khatoon, Jannath Bi

b) Names with father/husband as

KHURSHIDA BEGUM (Khan)

NASEEMA (Aziz)

c) Women names with Takhallus and like to known by thus name as

**in the western style as:**

ALA (Aziz Jehan)

KHATOON (A.R.)

d) **Women names with their husbands name should be mentioned as:**

ABDUL QADIR (Mrs.)

AHMED ALI (Begum)

For a detailed and comprehensive approach to Muslim naming pattern and cataloguing , please refer to the Book " Cataloguing of Indian Muslim Names by Mohammed Haroon".

## 3.    AUTOMATION OF URDU LANGUAGE COLLECTION

Since the library application softwares are by default in Roman script i.e. English language, the technical processing of Urdu language collection in an automated environment requires detailed study and planning before implementation.

Prior policy planning and decision making regarding the methodology to be adopted for conversion of existing database and simultaneously the current cataloguing, classification system is very essential, while initiating the automation process of Urdu language books.

For example decision regarding whether to install a separate software (In-house developed or commercial) to process Urdu language books or transliteration method could be followed to enter the bibliographical details in the system will be a major point to be resolved.

### 3.1    Constraints Encountered During Atomated Process Of Urdu Books:

A). *Bi-directional script problem:* Arabic ,Hebrew and Urdu are written from right to left, whereas the default direction of any standard code such as MARC, UNIMARC  etc. is from left to right.

B). *Sorting and indexing problem:*  Since the issue of sorting of multi-script data is still to be resolved and most of the existing automated indexing systems are by default for roman based scripts, the accessibility of language records through OPAC is a major constraint.

**C).** *Lack of precision:* **Due to lack of precision during input, users encounter difficulty in accessing/retrieving data in urdu language.**

**D).** *Lack of standardization:* **No standard spelling system is available for urdu language. Hence, no uniformity in rendering bibliographical data.**

**E).** *Assignment of key words/subject headings:* **Standardized codes like Sear's List of Subject Headings or Library of Congress Subject Headings are not available for assigning subject headings for urdu language collection.**

**F).** *The problem of data import/export:* **For books in urdu language (Arabic script), most of the software have no facility to import and export bibliographic data.**

**G).** *The Computer languages:* **The library application software for Indian languages are still in developmental stage.**

**3.2 Methods of Bibliographical Data Input**

**3.2.1 Transliteration Method: One way to overcome the constraint is to use the transliteration method to enter the bibliographical records of Urdu books into the system. In this method, the Urdu titles are entered using the Roman script.**

**The disadvantage in this is that, the users who are not familiar with the Roman script will not be able to utilize the facility to their advantage. The other problem is that there are chances of loosing semantic value (meaning) during the transliteration process. Phonetic deviation may occur due to the problem of script difference and lack of standardization of spelling/usage while in-putting the data. Further, due to lack of proper grip on the English language users may not be able to utilize the OPAC system. All these constraints can reflect on the quality of the database.**

**3.2.2 Entering Bibliographical Data Using Software: The most efficient method to input the bibliographical details is in the original script. This could be possible only if the software has such a facility. Very few such software exist in the Library and Information Science field. VIRTUA (VTLS) is one of the software developed exclusively for Library and Information Science field, and is increasingly becoming popular globally. This software has facility to enter bibliographical data in most of the major world languages (through its Language Editors). It also has taken care of the feature peculiar to Arabic languages i.e. right to left direction of writing.**

## 3.3     Urdu Software

**a). Urdu Software Digest**

**www.gy.com/www/ur.htm. 5K**

 **b). Urdu Software**

**www.axiscomputers.com/- 19k**

**c). Multilingual, Pashto, Farsi and Urdu software.**

**www.liwal.com/- 15k**

**d). Pakistan Database Management Services**

**www.pakdata.com/Urdu 98.php – 7k**

**e).Absolutely compatible urdu software**

**www.mshel.com/software.html - 22k**

**f). Urdu software, fonts etc.**

**www.urdustam.com/download/ - 8k**

## 4.     INDIRA GANDHI MEMORIAL LIBRARY AND ITS URDU DATABASE

The Indira Gandhi Memorial Library (IGML) is a central facility supporting the teaching and the research needs of the Hyderabad Central University.  Its users comprise of students, faculty and the non-teaching staff.

The library has a very rich collection of various types of reading materials in many languages. Currently, the library has 2.05 lakhs books apart from thesis and dissertations (about 2000) and 7,200 back volumes of journals.  The library subscribes to nearly 712 journals and a good number of magazines and Newspapers in several Indian languages.

Out of the total collection, approximately 21,000 books belongs to Urdu language

and the library subscribes to about 100 journals in Urdu language. The Library's collection  also includes books on other Indian languages such as Telugu, Hindi, Tamil, Kannada, Malayalam, Marthi, Bengali, etc.

The library purchases books and journals based on the recommendations received from the Dean/ Head of the Department of the concerned departments, within the limits of the budgetary allocations made by the Library committee to each department.

## 5.      DATABASE OF URDU LANGUAGE COLLECTION

The various efforts made by the Indira Gandhi Memorial Library (IGML) towards the technical processing and organization of the Urdu collection of the library and the various problems encountered and how they were overcome are described.

The collection development of Urdu books commenced along with the general books in English language in 1979. The books belonging to Urdu language are purchased based on the recommendations of the Head of the Department of Urdu Department (School of Humanities).

The various problems encountered and the adaptations done to overcome the constraints during the technical processing i.e. cataloguing, classification etc. are enumerated.

## 5.1     The Cataloguing

The cataloguing was done manually until automation was initiated.  In the absence of the availability of any standardized cataloguing code for Persian/Arabic languages, the AACR-2 code was adopted.  The bibliographical details of the Urdu language books were hand written in Urdu script (of various fields) similar to the English language books.

The various entries for each book i.e. the Main entry (Author entry) and all other added entries were all hand written which involved a lot of manual hours. The sorting and filing of cards in the catalogue cabinets also was a very cumbersome process.  Only the staff having the Urdu language expertise were able to handle these works.  It meant that the library had to have a person with a  dual knowledge of Library Science and Urdu Language.

After the automation was introduced the Shelf List card is being hand written and the bibliographical details of Urdu books are being entered in Roman script using transliteration method.

## 5.2    The Classification

The classification of Urdu language books has two pronged problems.  One is concerned with the language and literature class and the other concerned with the books of all other subjects.  The IGM Library uses the Dewey Decimal Classification (DDC) 18th edition to classify the books.

·        The books on Language and Literature Class:  The DDC  bias in the case of non-western languages is a major constrain while processing books in Indian Languages.  The numbers 490 and 890 "Other Languages" and other literature are used to club all the languages and literatures respectively (other than western origin). Thus all the Indian Languages have to share a small slot/subdivision alone with  not only the  languages of other Aryan and Dravidian languages, but among themselves.  Further, there is no scope to accommodate the innumerable Indian Languages and their dialects, which have not been enumerated in the classification scheme.

·        The Urdu Language: It belongs to the Indo-Aryan group and the number allotted in DDC classification scheme is 491.1 and 891.1 for language and literature respectively.  For Urdu language and urdu literature the number allotted is 491.439 and 891.439 respectively. Here again there is no scope and provision for accommodating dialects or any other aspects of urdu language as is possible in the case of western languages.

## 5.3    The Adaptation

To overcome such constraints, the IGM Library decided to adapt the system designed by the Central Institute of Indian Languages, Mysore called "Classification for Linguistics and Languages: a modified approach based on Decimal Classification"   (Mysore, 1973) with suitable modifications in the case of literature class.

The literature Class: The number for main literature class has been taken as base number with appropriate alphabetic index as an affix to individualise the books in Urdu language.  The sub-divisions are also adapted to construct number for various aspects like form number, dictionary, etc.

·        Example

| English Literature | | Urdu Literature | |
|---|---|---|---|
| | Poetry    821 | | U81 |

| | | |
|---|---|---|
| Drama | 822 | U82 |
| Fiction | 823 | U83 |
| Criticism-Poetry | 821D | U81D |

·          **Books Belongs To all other Subjects and Language:**

The classification number for books written in Urdu language belonging to all other subjects (Sciences, Social Sciences, Arts etc) and Urdu linguistics are same as that of books written in English language (Roman Script), with the individualizing alphabetic affix of the concerned language.

·        **Examples:**

**Roman script**                    **Urdu script**

| | | |
|---|---|---|
| Sociology | 301 | U301 |
| Chemistry | 540 | U540 |
| Agriculture | 630 | U630 |

## 5.4     The Subject Headings/Key Words

Standardized Subject Headings / Key words are not available as yet for urdu  language.  To overcome this lacunae the IGM Library provides the Key Words in English Language itself (and not transliterated), on the lines of those assigned in case of books in English Language.

## 5.5     The Book Number

The first three alphabets of the Surname of the author are used as the Book Number (written in Roman script) to simplify and facilitate quick shelving / retrieval.

·        **Examples**

|        |     |
|--------|-----|
| GHALIB | GHA |
| IQBAL  | IQB |

## 5.6     The Shelf Arrangement

The books in Urdu language are shelved as a separate collection.  The shelf organization and arrangement  in each shelf is similar to the English Language (Roman Script) books.  That is, the books are organized in DDC class number order 000 to 999, with books shelved from left to right in each rack.  The "U" alphabetic affix invidualising the Urdu books is ignored during shelving (as the books are kept as a separate collection).

## 6.     THE AUTOMATED TECHNICAL PROCESSING

The automation of the In-house operations of the IGM Library was initiated in 1989.The database of all the books written in English (Roman Script) is now available in Machine Readable form.  The books are processed through the LIBSYS (LIBSYS Corporation, Delhi ) library application software. LIBSYS system is an integrated software and has six Menus – Acquisition, Cataloguing, Circulation, Serial Control, Article Indexing and OPAC (On-line Public Access Catalogue).  Apart from these there are three modules for House Keeping Operations – Record Keeping, System set-up and Utilities.

## 6.1     The Automated Retrospective Conversion

The retrospective conversion of Urdu database is almost at the completion stage.  The procedure of cataloguing and classification etc. are as described earlier, with slight difference from the manual atmosphere, which has been explained under the Current conversion.

## 6.2     The Current Conversion

The books ordered and received from the book suppliers are entered in the Acquisition Module through transliteration method.  Once the Acquisition process is complete, the data is available in the Technical Module.  The procedures of cataloguing and classification have been already discussed.  The difference is that only one manual card is prepared, which is maintained as a Shelf List Card.  As once the bibliographical details are entered in the system, the software is programmed to generate the various Access Points - Author, Title, Classified Number, subject, Words in Title and combination search (which are similar to the added entry cards).

The Bibliographical details are entered through the Transliteration Method.  The Key Words are assigned based on the DDC Schedule's Subject Headings and Clue Words in the book itself and are entered in the English Language itself (not transliterated).  In the automated environment, the unnecessary noise words like prepositions, conjunctions etc are all eliminated and hyphens are entered between key words.  This facilitates efficient retrieval of information as each key word/key words set is treated as a separate entity by the system.

· **Example:**

**SUBJECT HEADINGS :**         CULTURE OF INDIA

                                                           HISTORY AND POLITICS OF INDIA

COMPUTER APPLICATIONS TO THE SUBJECT PHYSICS

          **KEY WORDS :**                       CULTURE -- INDIA

                                    I           NDIA  -- HISTORY  --  POLITICS

                                                   COMPUTER APPLICATIONS  --  PHYSICS

The IGM Library is now planning to digitalise selected core collection, which will include books in Indian language too.  To facilitate this process, the library has installed the VTLS software system, which has the advantage of allowing the entry of any Indian Language in its own script.  Currently the detail of the books in the Urdu collection of the IGM Library's database is accessible through the LAN. Hopefully, soon it will be possible to browse the content itself through the Internet.

## 7.    CONCLUSION

To overcome most of the problems enumerated in the paper , suitable classification  scheme, cataloguing system, which are amenable to easy synthesis and shelving require to be adapted.  During planning and preparation of an adapted system of classification and cataloguing, care should be taken to make it staff and user friendly. Further, when embarking on automation process, it is necessary to make sure that the Application Software selected has facility to systematically process the books in Indian languages  in their original script. It is hoped that this paper will act as a catalyst in making the Library and Information Science professionals dealing with Indian Language database to think seriously about the constraints and the existing lacunae and pressurize the software designers, producers to design suitable Application Software system for efficient processing and organization of Indian Language Databases.

# REFERENCE

1.	Biligiri, H.S. and C.R.Sulochana: Classification for Linguistics and Languages: a modified approach based on Decimal Classification. Mysore, Central Institute of Indian Languages, 1973.

2.	Gangu, B.T. and Yakub Ali: Hindi Database Management. Paper presented at 47th ILA conference, Warangal.

3.	Haroon, Mohammed:  Cataloguing of Indian muslim names.  Delhi, Indian Bibliographies Bureau, 1984.

4.	Natarajan, M. : User's attitude towards Electronic Journals: comparison and evaluation.  Paper presented at NACLIN 2001, Hyderabad.

5.	Ramachaudhuri, V., V.J.Susheela and V.Uma ; Need for multilingual database in Indian library and Information Centres. CALIBER – 1996.

6.	Sehagal, R.L.: Handbook of Library Software Package. New Delhi, ESS ESS Publications, 1998

7.	www.google.com/urdu software.

## BRIEF BIOGRAPHY OF AUTHORS



*Dr.  (Mrs) B.T.Gangu* is currently working at the Indira Gandhi Memorial Library, University of Hyderabad, as an Assistant Librarian.  She obtained her doctoral degree in Computer and Information Science.  Her areas of expertise include library automation and classification.



*Mr. Yakub Ali*  has M.A. (AIHCA) and M.L.I.Sc. form University of Madras, is currently working at Indira Gandhi Memorial Library, University of Hyderabad.  He has more than 15 years of professional experience and specializes in processing the books in Indian Languages.  He is a member of professional bodies in the field of Library and Information Science.  He has credit the publication of some articles.