March 2022

ISSUES AND CHALLENGES IN INDIAN MULTI-LINGUAL AND MULTI SCRIPTS BIBLIOGRAPHIC RETRIEVAL SYSTEMS

Shivpal Singh Kushwah Dr.
Rani Lakshmi Bai Central Agricultural University, Jhansi, shivpal_s@yahoo.com

Ritu Singh Dr.
Bundelkhand University, Jhansi, rtkushwah@gmail.com

ISSUES AND CHALLENGES IN INDIAN MULTI-LINGUAL AND MULTI SCRIPTS BIBLIOGRAPHIC RETRIEVAL SYSTEMS

Dr S S Kushwah
University Librarian
Rani Lakshmi Bai Central Agricultural University, Gwalior Road Jhansi, Uttar Pradesh, India

Dr Ritu Singh
Assistant Professor, Library & Information Science
Bundelkhand University, Kanpur Road, Jhansi, Uttar Pradesh, India

Absract

Multilingual retrival systems are very important for countries like India, where we have multiple scripts and many languages for verbal and written communication. The creation of a multilingual interface for the retrival, management, and processing of information and knowledge needs systematic efforts and requires related features in the Library Management Software and the availability of bibliographic records in catalogues.

This paper is an effort to study various related aspects of multilingual record creation and retrival provisions in Union Datanases. The development of a multilingual environment for accessing and retrieving library resources among the users as well as library professionals is essential. The article is divided into five sections. The first section deals with the introduction. It covers the multilingual system, with reference to multiscript bibliographic control, the objective, and methodology of the study. The second section examines the contours of multilingual bibliographic control, in particular in the Indian context. The third section deals with the multilingual cataloguing procedure and The fourth section deals with the "Situation in India in Terms of Multilingual Bibliographic Database and Issues." issues, challenges, and solutions. The fifth and final section of the article is concerned with the conclusion.

The objective of this paper is to analyse the issues and challenges in the Indian multi-lingual and multi-script bibliographic retrieval systems. Through the analysis of available bibliographic data and database interfaces, the authors evaluate the adoption of multilingual and multiscript processes and procedures for bibliographic data creation.

The study's findings will aid in understanding the current status of multilingual bibliographic record creation and the need for policy-level intervention to maintain and develop multilingual records for the creation of qualitative bibliographic databases.

Keywords: Union Catlogue, Multi Scripts Bibliographic Retrieval Systems, Indian Multilingual Retrival Inititives.

INTRODUCTION

It is always necessary for libraries to develop an interface for its users to retrieve information effectively and in the easiest possible manner. Therefore, libraries are cataloguing their resources, developing databases and acquiring online and offline resources. Language is one of the important mediums for the written, verbal communication and dissemination of information. Therefore languages and scripts have emerged to facilitate the

communication of thoughts, information, knowledge and ideas. The publishing industry, authors as content creator, and scholarly societies as content creator and users are developing content and producing scholastic output in various languages and scripts. Therefore, the libraries of the world have to work in more than one language to serve their users community/clienteles. The collection in the libraries includes books and other documents published printed in different languages written in defferent scripts due to users demand. Libraries acquires books in local and regional languages, in addition to books written in the national language (Rajbhasha) and in English. It is universally known that a collection of the libraries is of no use, has no value, unless until it is utilised by the users. Therefore, an effective and useful catalogue comes into the picture. It is an essential tool for the users' community to facilitate retrieval of reading materials available in the library.

In the case of India, which is a country of different languages and cultures, many scripts and languages are in practice, for the communication of thoughts and ideas. The Indian libraries are acquiring printed and published literature in different languages. As per Articles 344 (1) and 351 of the Indian Constitution, the eighth schedule includes and recognises 22 official regional languages, **i.e.** Assamese, Bengali, Bodo, Dogri, Gujarati, Hindi, Kannada, Kashmiri, Konkani, Maithili, Malayalam, Meitei (Manipuri), Marathi, Nepali, Odia, Punjabi, Sanskrit, Santhali, Sindhi, Tamil, Telugu, and Urdu (Mandal, 2018; Rao, 2021; Sardar, n.d.). The ancient knowledge originated in Indian languages, scripts and its primary sources, philosophical thinking available in original local language. Indian ancient science and philosophy originated from india is also available in Indian languages.

Libraries are developing catalogues, indexes, and bibliographical systems to cater the requirement of its users. A catalogue is a list or index of reading materials (i.e., books, journals, and other types of reading material) available in the library/librariesy. Cataloguing standard provides a framework for maintaining uniformity and consistency in the cataloguing and bibliographic product development processes (Chandrappa & Harinarayana, 2018). Multilingual computing is based on the encoding of the language characters, initially, the cataloguing was started with the english language only, the reson behind this was the availability and adoption of enlish language for computation. As we all know, the larger parts of the world communicate with each other either in their local language or using other mediatory languages such as english (*Multilingual Information Management:*, n.d.). The libraries managing multilingual content are facing a big challenge while disseminating library services to the multilingual user community. The quality of the records plays a very important role. The concept of quality in the case of bibiliographic products varies with reference to multi-script multilingual databases. In this case, one of the most important quality indicators is the quality of the record in terms of rendering, compliance with AACR Rules, unique, error-free entry, proper transliteration, use of preferred data fields, and so on.In the present case, products mean a bibliographic retrieval interface, a bibliographic database, bibliographies, etc. It was developed for the user community to retrieve books from the library. Consistency in cataloguing, especially in multilingual cataloguing, has its own importance for facilitating library users in an effective manner.

Objective and scope of the study

The multilingual platform and multilingual bibliographic content creation is important, similarly consistency in the process, procedure and in data is also important. In this study, an attempt is made to know the procedural approach adopted for multilingual records creation in various national initiatives with following objectives:

1. Find out the multilingual record creation approach adopted by the Indian libraries.
2. Find out issues and challenges related to multilingual bibliographic record creation.
3. Study and observe the available records in two mazor union catalogues (i.e. Indian Digital Ensemble of Agricultural Libraries (IDEAL) of ICAR and IndCat of UGC), and Union Catalogue hosted on Indian Culture Portal as part of the National Virtual Library of India project, funded by the Ministry of Culture, Government of India evaluated and reviewed to find out the adopted approach for multilingual record creation.
4. Study the existing situation of Multilingual Bibliographic Record Creation and Retrival Approch.
5. Make recommendations/suggestions for improvement Multilingual Bibliographic Database and Retrival System.

Methodology

The study is a result of a critical analysis of the adoption of multilingual multiscripts bibliographic record creation standards in Library Management Softwares and cataloguing practices related to multilingual bibliographic databases and language bibliographic record creation. The context of the study is with special reference to Indian languages. The observations drawn from the from the bibliographic data contributed by the various academic and research libraries in Union Catalogues. The research question that the articles seek to find an answer is: what are the challenges that have come with language bibliographic record creation and how are libraries responding to it? In order to answer these questions, a number of methods have been used, which are based on inputs from primary and secondary sources. Theoretical, analytical, and descriptive methods have been used to test the hypothesis by examining and referring to available literature, data, and insights from secondary sources.

CONTOURS OF MULTILINGUAL BIBLIOGRAPHIC CONTROL

It is common for the libraries to have bibliographic data in more than one language; transcribed descriptive data in the *language(s) of the item(i.e. books/Journals etc)*; with access points and notes in the *language of the catalog*. For providing alternative access options, the libraries are also providing access points and notes in more than one languages. The technological developments allow the use of multilingual and muulti scripts in catalogues. However, a significant challenge for libraries is the lack of established processes and procedures, as well as the lack of practical guidelines adoption (MacDougall 1997). The retrieval of the information from the database to display in a different language, process data, and add new data carries many issues and challenges. The libraries may creates records in multiple languages in a database using DBMS/RDBMS interface can be managed through software (Mandal, 2018). Earlier catalogers were creating fully transliterate bibliographic records for non-Latin script collections/ documents; at international level the "Vernacular" MARC cataloging for JACKPHY (Japanese, Arabic, Chinese, Korean, Persian, Hebrew, & Yiddish) began in 1982; ALA-LC Romanization Tables. Transliteration schemes for non-roman scripts. Developed and maintained by Library of Congress. Washington, D.C.: Cataloging Distribution Service, Library of Congress, 1997. - For entering the records in roman letter • Indian Standards - Indian script code for information interchange -ISCII. New Delhi: Bureau of Indian Standards, 1991. (IS 13194: 1991) ANNEX-F: ROMAN SCRIPT TRANSLITERATION (The National Library at Calcutta standardized the diacritic marks to

be used for romanization of Indian scripts, in 1988 (“The National Library Newsletter”, June 1988) - For entering the records in roman letter (INFLIBNET, n.d.).

The MARC record creation was based on the "dual script" concept. In the old "MARC-8" (8-bit) environment, the computing was based on ASCII 8 bit character coding. In 1990s Unicode was adopted, major MARC systems implemented Unicode after January 2000; OCLC and RLIN implemented Unicode by 2004; the Library of Congress implemented Unicode in 2005. In india ISCII was adopted in the early stages of multilingual computerisation with support of the of products developed by CDAC named as Graphic and Intelligence based Script Technology (GIST) Cards. The GIST card has been adopted for the computerization in India, such as the Land Records Program, the Election Commission Identity Cards, and citizen surveys etc. This technology was adopted by INFLIBNET for the development of a integrated library automation system SOUL (software for university libraries). Chadrakar (2004) briefly reviewed the introduction of Unicode technology in India and its implementation in SOUL software. With the development of UNICODE the LMS has also started adopted it in their interface for data entry in case of catalogue module and retrival in case of OPAC. Many library library LMS such as SOUL, Libsys, VTLS, KOHA are in practice in Indian libraries, due to budget crunch the libraries are moving towards open sources (Gautam et al., 2008). The software named SOUL developed by INFLIBNET supports many fetures but this also has closed template for data entry it supports Unicode based multilingual catalogue features, but bilingual or multilingual record can not be created in the software with provision of 880 tag “alternate graphic representation” as suggested in MARC 21 “Model A”. Although the SOUL can be used for multilingual cataloguing purposes with UNICODE (Vaghela et al., 2008).

Open Source Software which are adhered to international standards such as MARC, Unicode, and ISO, library cataloguing like KOHA are providing option to create multilingual multiscript records using option Model A or Model B.The cataloguer is required for the development of a multilingual collection in any library, which requires additional knowledge of multilingual languages and scripts, transliteration, and diversity knowledge of different languages collections.

The language bibliographic control consists of encoding standards, methods and process of multilingual bibliographic record creation and retrival system. In India, various languages are spoken by the population at large. The computerized system process data at their backend interface in the form of numbers and at front end, it shows letters and other characters by assigning a encoded value in form of number (digits) to each one. Prior to the development of UNICODE, there were many different systems related to language character encodings, for assigning encoded value (numbers) to the characters was in practice. Due to the late emergence of UNICODE-based solutions for language computing, most libraries of national repute have created their bibliographic data mainly in English. A *character set is very important for* written communication in particular languages. It pla ys a very important role in a writing system. A writing system is a method of visually representing verbal communication, based on a script and a set of rules regulating its use. Most languages have a single character set, and similar character sets are often used by a number of languages (e.g., variants of the Roman alphabet are used to write English, Spanish, Finnish, Dutch, etc.). The Devangar script is used to write the Sanskrit, Prākrit, Hindi, Marathi, and Nepali languages.

The emerging technology relating to language content creation has made it possible for computer systems developers to develop language compliance tools and applications that can understand and synthesise spoken and written human languages. Many tools and

applications related to speech processing (recognition, understanding, and synthesis), information extraction, handwriting recognition, machine translation, text summarization, and language generation are emerging. In this journey encoding standards related to character encoding are playing a very important role.

ASCII: The English language writing is based on a **script** consisting of Latin script, the script used for writing the English language (set of letters), and an 8-bit character coding which allows coding of characters (2 X 2 X 2 X 2 X 2 X 2 X 2 X 2 = 256). The lower part of the language characters is defined as per standard IS10315:1982 (ISO 646 IRV), it is a 7 bit coded character set for information interchange, known as ASCII, which stands for American Standard Code for Information Interchange character set (*ASCII Codes - Table of Ascii Characters and Symbols*, n.d.)

ISCII : The Indian Script Code for Information Interchange (ISCII) is based on an 8 bit character code which allows coding of characters (2 X 2 X 2 X 2 X 2 X 2 X 2 X 2 = 256) for use in computers and communication media. The top 128 characters meet the requirements of all the Indian scripts based on the ancient Brahmi script. The remaining 128 are accounted for by the **script,** consisting of the Latin script, the script used for writing the English language as defined in ASCII (*ISCII (Indian Standard Code for Information Interchange)*, n.d.). There are 22 officially recognised languages in India. For written communication, they are based on scripts. Apart from Perso-Arabic scripts, all the other 10 scripts used for writing Indian languages are originated from the ancient brahmi script. They have common phonetic structure, therefore they are making a common character set. The ISCII Code table codified all the characters required in the brahmi based Indian scripts.

UNICODE: Unicode Standard has three encoding systems i.e. UTF8, UTF16 and UTF32, which is being used internationally by the software and computing industry for the development of Multilingual Softwares and applications. Unicode has made our life simpler in terms of communication in local language. UNICODE standards assigns each character a unique numeric value and name. The total **159** scripts of the languages of world are supported by the Unicode version **14.0** (2021) Standard include all of those listed in the following table (The Unicode Consortium, n.d.).

MULTILINGUAL CATALOGUING PROCEDURE

Computerised cataloguing is depends upon the guidelines and procedures developed for the cataloguing nd their adoption by the software developer and cataloguing professionals. As we all know practice is repetition of an activity / procedure for performing a task, appropriate practise is essential for language bibliographic data development. The creation of computerised bibliographic records the bibliographical standard MARC 21 has suggested two models for the creation of multi lingual records, as mentioned below.

Model A: Translitaration (based on *vernacular and transliteration), a* vernacular**, or** vernacular language, is a term used to refer to a local language or dialect, as distinct from what is seen as a standard language. The regular fields of the record may contain data in different scripts and in the vernacular or transliteration format. MARC 21 field 880 can be used when data needs to be duplicated to be expressed

in *both* languages, i.e., in the original vernacular script and transliterated into one or more scripts. The detail of the field used in a multilingual record using Model A option.

100 (author)
245 (title statement)
246 (varying form of title)
500 (general note)
650 (subject added entry – topical term)
700 (added entry – personal name)
250 (edition statement (r))
260 (publication, distribution, etc.)
300 (physical description)
880 (alternate graphic representation)

Sample Record Created using Model A approach (MARC view)

000 -LEADER
fixed length control field 00597nam a22001817a 4500
003 - CONTROL NUMBER IDENTIFIER
control field OSt
008 - FIXED-LENGTH DATA ELEMENTS--GENERAL INFORMATION
fixed length control field 210223b ||||| |||| 00| 0 eng d
020 ## - INTERNATIONAL STANDARD BOOK NUMBER
International Standard Book Number 9788189242114
040 ## - CATALOGING SOURCE
Transcribing agency rlbcau
100 ## - MAIN ENTRY--PERSONAL NAME
Linkage 880-01
Personal name Premchand
9 (RLIN) 1384
245 ## - TITLE STATEMENT
Title Godan
Remainder of title / by Munshi Premchand
260 ## - PUBLICATION, DISTRIBUTION, ETC.
Linkage 880-01
Place of publication, distribution, etc. Delhi
Name of publisher, distributor, etc. Archna Publication
Date of publication, distribution, etc. 2019
300 ## - PHYSICAL DESCRIPTION
Extent 336 P.
942 ## - ADDED ENTRY ELEMENTS (KOHA)
Source of classification or shelving scheme
880 ## - ALTERNATE GRAPHIC REPRESENTATION
Linkage 100-01
-- प्रेमचंद
880 ## - ALTERNATE GRAPHIC REPRESENTATION
Linkage 245-02
-- गोदान / मुंशी प्रेमचंद
880 ## - ALTERNATE GRAPHIC REPRESENTATION
Linkage 260-03

-- दिल्ली
-- अर्चना पब्लिकैशन
-- 2019

The abovementioned examples show that the record contains bibliographic data in both scripts. In this example, the language of cataloguing is English Sript (Cyrillic script), and the bibliographic item is a mixed Russian language (Cyrillic script) and English language (Latin script) text. This formate is consistent with the cataloguing practise for fully transliterated scripts in Latin script catalogues; it avoids the issue of repeating non-repeatable fields, such as field 100 or field 245. It is easier to suppress non-Latin data in Latin script cataloguing environments; clear distinctions are made between data in different scripts (vernacular & non-vernacular) (Barry, n.d.).

Model B: *simple multiscrip, bibliographic records;* all data is contained in regular fields; script varies depending on the requirements of the data. Although the Model B record may contain transliterated data. As per MARC 21 format, the most commonly used tags used for simple cataloguing are ; in case of multilingual cataloguing also libraries are using same tage for entering data of different language. The detail of the field used in a multilingual record using Model A option. The detail of the field used in a multilingual record using Model B option.

100 (author)

245 (title statement)

246 (varying form of title)

500 (general note)

650 (subject added entry – topical term)

700 (added entry – personal name)

250 (edition statement (r))

260 (publication, distribution, etc.)

300 (physical description)

650 (subject added entry--topical term)

700 (added entry--personal name)

Sample Record Created using Model B approach (MARC view)

000	00498nam a2200169Ia 4500
008	140201s9999 xx 000 0 und d
020	_a9789381450154
100	_aकुमार, संजीव
245	_aसमेकित कृषि प्रणाली _b
250	_a1st
260	_bNew India Publishing Agency

_aNew Delhi
_c2012

300 _a331
650 _aHindi Books
700 _aसिह, सती शंकर
700 _aशिवानी

The alternative (typically Latin) script is treated similarly to the original (sometimes non-Latin) script. There is no need to deal with any particular field for embedding other script; At the field and subfield level, scripts may be blended considerably more successfully. The creation and maintenance of records is substantially easier since there are no links between fields **(Barry, n.d.)**.

THE SITUATION IN INDIA IN TERMS OF MULTILINGUAL BIBLIOGRAPHIC DATABASE AND ISSUES

Libraries and the publishing industry, as well as national organisations such as national libraries, were putting forth efforts to develop products in form of union catalogues and national bibliographies. The Computerised bibliographic record creation started with the development of library automation systems, the development of union catalogues and national bibliographic control initiatives. In terms of India various cataloguing and bibliographical initiatives have been made by the organisations involved in the development of the Union Catalogues.The union databases or virtual union catalogue of the library resources are one of the important services of the libraries and consortiums. In India, at national level there are two mazor organizations INFLIBNET, UGC (University Grants Commission), and CeRA ICAR (Indian Council of Agricultural Research) developing databases of resources available in the libraries comes under their domain. Bibliographic catalogues play an important role in bibliographic control in terms of the bibliographic information availability, helps in locating required books in the library, also help in avoiding the effort of creating duplicate catalogues of records by providing copy catalauging option. Another purpose of the bibliographic databases is to provide or distribute information to the users using single window access about the availability of documents in the libraries. It is a useful tool that can be accessed from the desktop 24x7.

UGC INITIATIVES

INFLIBNET maintains and develops union databases for various libraries named as IndCat. This activity was started in the year 1991. Currently, it consists of 1,24,77,860 unique titles with 1,67,94,578 holding details of 203 **(*IndCat,* n.d.)**. The data contributed by the participating universities is pooled together hosted on central cerver at INFLIBNET, Gandhinagart. The process of the database development of the IndCat include the receiving of bibliographic records from the university libraries in MARC or MARCXML format. The libraries are creating records in using library automation system such as SOUL, KOHA, Libsys, VTLS, etc. At the time of receiving data the cetre is checking and validating data, Importing validated data as MARC File and converting it in CSV file format. The process of the union database development completed through UCMS . In bigining of the database development process the Centre has used CDS/ISIS (Computerised Documentation Service / Integrated Set of Information Systems). The UCMS is windows-based application, developed in-house by the Centre. The UCMS is used for authentication and re-formatting of bibliographic records received from universities as well as for merging new records into the

existing union catalogue. The database provides interface for copy cataloguing using newly developed OCS interface **(*Online Copy-Catalogue System*, n.d.)**. ISO 23950:1998 Information and documentation - Information retrieval (Z39.50) is a internationally accepted protocol specification and standard for copy cataloguing, this standard was last reviewed and confirmed in year 2020 to keep the version remains current **(ISO 23950:1998, n.d.)** but INFLIBNET has not adopted z39.50 for its IndCat database.

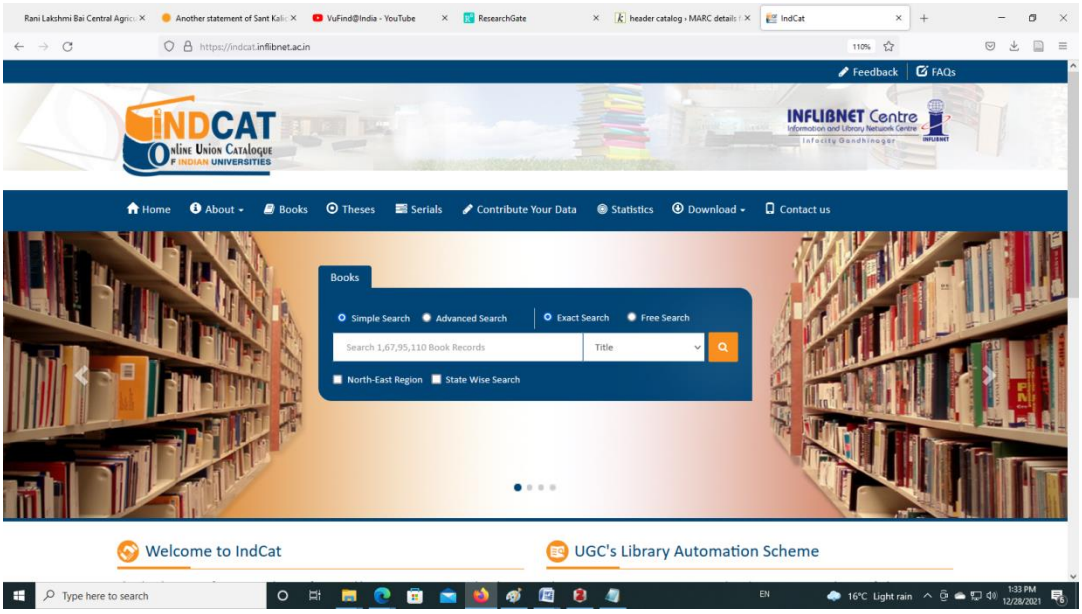

Fig.1 INFLIBNET Ind Cat

Normal View | MARC View

Tag	Ind	Sub-Field	Value
000			00629amm a0000169ua 4500
001		a	12904177
003	__		IN-AhILN
008			120126n ii |||gr||||Z|||||eng||
035		a	(BHUL_221005)33301
082	\	a	0152,1:g(Y)
100	00	a	विजयवर्गीय, प्रेमचन्द
245	00	a	आधुनिक हिन्दी कविता का सामाजिक दर्शन
245	00	h	Textual Documents
260	00	a	Jaipur, India
260	00	b	Babhana
260	00	c	1977
300	00	a	421p
650	00	a	हिन्दी साहित्य - सामाजिक दर्शन
904	00	a	UC
905	00	a	5-1-84
906	00	a	Hindi

Fig. 2 Sample Language Record from IndCat

Based on bibliographic data at the website it is found that it has:

Features

- Union catalogue has total 203 universities who has contributed their bibliographic records.
- The system has 1,24,77,860 unique titles.
- OCS is developed by the INFLIBNET. OCS is not standard way to retrieve / copy records. It is a Window based application, which communicates directly to IndCat. It is fully compatible with SOUL an automation software developed for automating Indian libraries.
- Database provides searching facility using different parameters i.e. Title, Author, Subject, Publisher, Year, ISBN etc. Facilitate browsing, copy and copy as a new records feature. Create new catalogue of book and save it into Library Software and IndCat Server. It supports multilingual feature.

Issues

- There is no serarch option to retrieve language record restricted search.
- ISO 23950:1998 Information and documentation — Information retrieval (Z39.50) last reviewed and confirmed in year 2020 not integrated with IndCat.
- OCS developed by INFLIBNET is not standard interoperable standard.
- IndCat has no interface for retrival of language record.
- Linked data field not used in any of the record available in IndCat.
- Data is hosted on central server, that restrict live updation.

ICAR Initiatives

ICAR has made efforts to upgrade, strengthen, and automate libraries of the academic and research institutions comes under its domain through various programmes (i.e. NATP, NAIP, and NAHEP). Union catalogue of agriculture AgriCat by integrating the libraries of National Agricultural Research System (NARS) institutions, State Agricultural Universities (SAUs) and ICAR institutions started under NAIP a project. Later on the free and open source software KOHA was adopted to automate the libraries. The cloude based database was developed by hosting multilibrary holdings on cloude named as IDEAL.

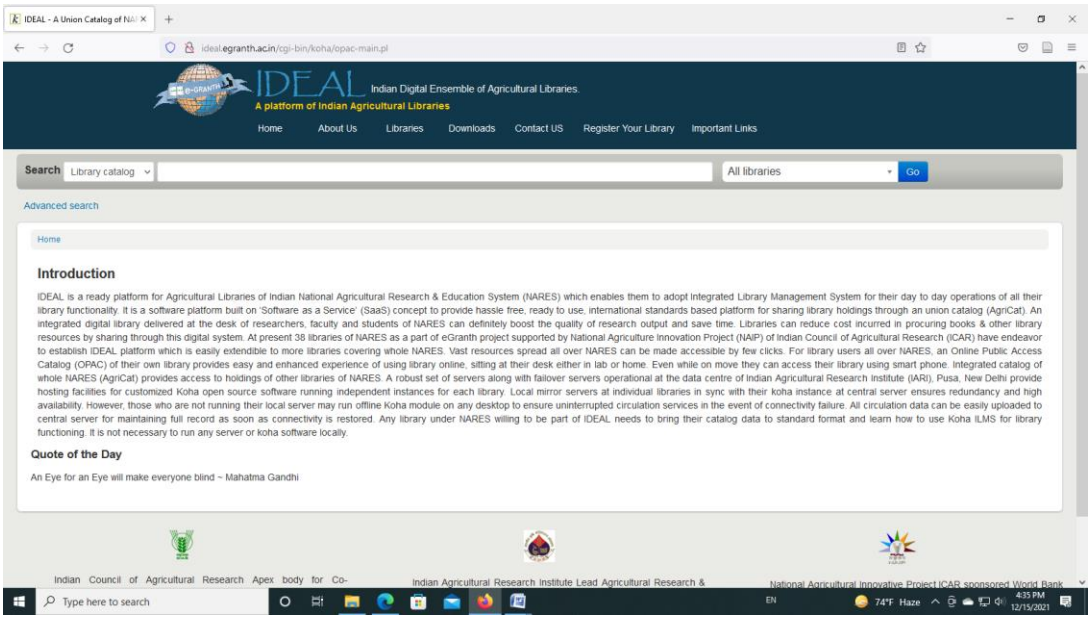

Fig.3 ICAR IDEAL

The libraries those were automated using other library LMS, their data converted and migrated from other LMS system to KOHA. The Online Union Catalog of Agricultural Libraries was initiated by ICAR under project NAHEP named "IDEAL." This catalogue contains bibliographic records contributed by the various agriculture universities (Kandpal & Rawat, n.d.).

Based on bibliographic data at the website of IDEAL it is found that it has:

Features

- Total 38 University and R& D libraries contributed their data.
- Earlier libraries were using defferent library automation system now they have migrated their data in KOHA.
- It has 4476 Record related to hindi books**.**
- Z39.50 Server option is available in KOHA but not configured in IDEAL system.
- Data is hosted on central server, it is not a right approach, it alwase provides old data.
- Many Libraries have traqnsliterated data in
- Data is hosted on central server, it is not a right approach, it alwase provides old data.
- Many Libraries have traqnsliterated data in
- It provides searching facility using different parameters i.e. Title, Author, Subject, Publisher, Year, ISBN etc.

Issues

- ISO 23950:1998 Information and documentation — Information retrieval (Z39.50) last reviewed and confirmed in year 2020 not integrated with IDEAL.
- Data is hosted on central server, that restrict live updation.
- Data created in the system are based on model A approch

The data available in the system shows that all the libraries have created records using Model B of the *simple multiscrip, bibliographic records;* all data is contained in regular fields; script varies depending on the requirements of the data. Although the Model B record may contain transliterated data. Model A is preferred if the same data is recorded in both the original vernacular script, field 880 linked data is uses.

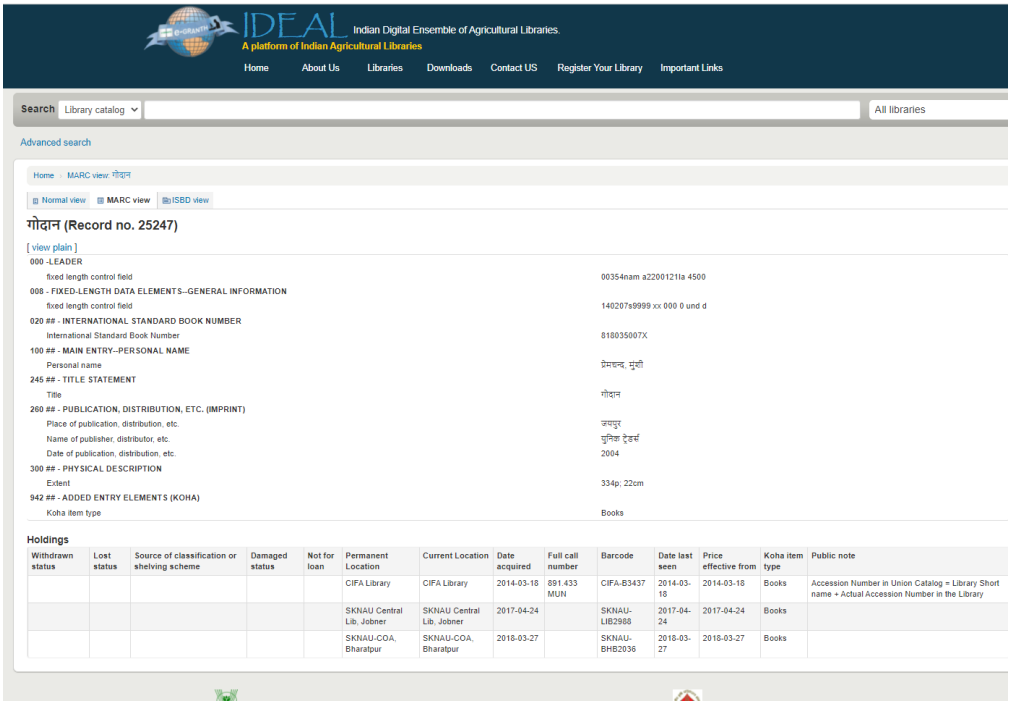

Fig.4 Multilingual Record in ICAR IDEAL

Union-Catalogue Ministry of Culture Initiatives

Recognizing the need of the users community a Union Catalogue is developed with partnership of by the Ministry of Culture, IIT Bombay, IGNOU. It is a a platform that hosts data of cultural relevance from various repositories and institutions all over India. The union catalogue is hosted using costomised application.(*Union-Catalogue | INDIAN CULTURE*, n.d.) it has included following libraries data :

1. Lalit Kala Akademi (2391)
2. Rampur Raza Library (9593)
3. Sangeet Natak Akademi (32900)
4. Raja Rammohun Roy Library Foundation (35270)
5. Connemara Public Library (112959)

6. National Archives of India (134109)
7. Indira Gandhi Rashtriya Manav Sangrahalaya (236341)
8. Central Secretariat Library (324727)

Based on bibliographic data at the website it is found that it has:

Features

- Earlier libraries were using defferent library automation system now they have migrated their data in KOHA.
- Data is hosted on central server, it is not a right approach, it alwase provides old data.
- Traqnsliterated romanised records are created for multilingual data.
- Data is hosted on central server, it is not a right approach, it alwase provides old data.
- It provides searching facility using different parameters i.e. Title, Author, Subject, Publisher, Year, ISBN etc.

Issues

- ISO 23950:1998 Information and documentation — Information retrieval (Z39.50) last reviewed and confirmed in year 2020 not integrated with IDEAL**.**
- Data is hosted on central server, that restrict live updation.
- Data created in the system are based on model A approch

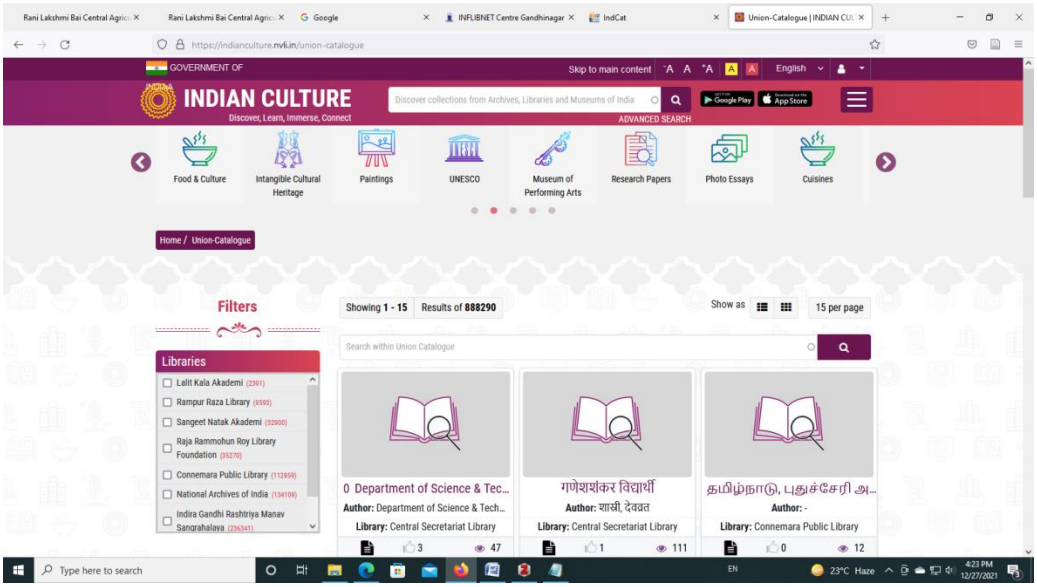

Fig.5 Union Catalogue Indian Culture

Based on observation of the records available in the above mentioned national bibliographic databases it is found that the databases has following issues related to multilingual cataloguing:

- It is a big issue for the databases that, the majority of participating libraries that are contributing data are creating transliterated records for their multilingual documents. This type of data does not provide the option to retrieve records using a language interface. Furthermore, there are lots of inconsistencies in transliterated bibliographic records. Therefore, the retrieval interface is not providing the desired results to the user community.

1. Based on a web survey of the catalogues, it is observed that most libraries have created transliterated data, but very few libraries have adopted an approach to creating multilingual records as per prescribed practice.
2. It is observed from the submitted data available in the mazor union catalogues (i.e., Indian Digital Ensemble of Agricultural Libraries (IDEAL) of ICAR and IndCat of UGC), and the Union Catalogue hosted on Indian Culture Portal as part of the National Virtual Library of India project, funded by the Ministry of Culture, Government of India, that data is lacking consistency in terms of policy and procedure for multilingual record creation.
3. The issues in the Multilingual Bibliographic Record Creation and Retrieval Approach can be addressed with proper guidelines and procedure for record creation. Users will benefit from properly created uniform consistent data in their retrival as well.

- Because Protocol Z39.50 was not integrated and implemented by the bibliographic agencies, libraries are not in a position to use the copy cataloguing option properly to maintain consistency.
- Indian properitory software does not provide an interface for linked data MARC 21 field 880 for alternate graphic representation.
- It is observed that the UNICODE is integrated into various library automation systems, but due to a lack of multilingual data entry guidelines, the records are inconsistent.
- Real-time virtual union catalogues developed using standard protocols such as "Integrated Library System–Discovery Interface (ILS–DI)" may be a better alternative to union catalogues created on a single server using KOHA LMS (IDEAL of ICAR and Vertual Union Catalog of Ministry of Culture Initiatives) and Ind Cat by INFLIBNET using its own proprietary tool, UCMS.
- All the efforts related to the union catalogue or multilingual union catalogue need to support interoperability and open standards.

CONCLUSION

In light of the above discussion, it can be said that multilingual bibliographic standards need to be adopted by libraries for qualitative record creation. National initiatives need to focus on quality content dissemination and promote consistency and uniformity in multilingual record creation. The National Library at Calcutta standardised the diacritic marks to be used for the Romanization of Indian scripts in 1988, but after the adoption of UNICODE, the national library has not adopted a policy for UNICODE based multilingual record creation. A common national guideline for multilingual bibliographic record creation is much needed for maintaining consistency and uniformity. Standards such as the Z39.50 protocol and the Integrated Library System–Discovery Interface (ILS–DI) protocol need to be incorporated into national union catalogues to provide a hessel-free, easy interface for searching.

Reference

Barry, R. K. B. R. K. (n.d.). *MULTIPLE LANGUAGES, MULTIPLE SCRIPTS, AND MULTIPLE HEADINGS IN MARC AUTHORITY RECORDS*.

Chandrappa, & Harinarayana, N. S. (2018). *What constitutes cataloguing quality? A review of issues and perceptions*. *1*, 85–93.

Gautam, J. N., Kushwah, S. S., & Singh, R. (2008). *Library Automation and Open Source Solutions Major Shifts & Practices: A Comparative Case Study of Library Automation Systems in India.*

IndCat. (n.d.). Retrieved December 15, 2021, from https://indcat.inflibnet.ac.in/

ISCII (Indian Standard Code for Information Interchange). (n.d.). Retrieved December 20, 2021, from https://www.cdac.in/index.aspx?id=mlc_gist_iscii

ISO 23950:1998. (n.d.). *ISO 23950:1998*. ISO. Retrieved December 16, 2021, from https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/02/74/27446.html

Mandal, S. (2018). Development of Multilingual Resource Management Mechanisms for Libraries. *Library Philosophy and Practice*, *2018*.

Multilingual Information Management: (n.d.). Retrieved December 13, 2021, from https://www.cs.cmu.edu/~ref/mlim/index.shtml

Online Copy-Catalogue System. (n.d.). Retrieved December 15, 2021, from https://ocs.inflibnet.ac.in/

Rao, G. (2021). Linguistic Diversity, Education, and Democracy in India. *Philology Sciences*. https://www.academia.edu/63423672/Linguistic_Diversity_Education_and_Democracy_in_India

Sardar, B. (n.d.). *Linguistic Diversity of India and Its Official Language*.

The Unicode Consortium. (n.d.). Unicode. *Unicode*. Retrieved December 20, 2021, from https://home.unicode.org/

Union-Catalogue | INDIAN CULTURE. (n.d.). Retrieved December 27, 2021, from https://indianculture.nvli.in/union-catalogue

Vaghela, D., Patel, Y., Rayka, D. L., & Chandrakar, R. (2008). *Unicode Based Multilingual Catalogue Module: A New Feature of SOUL*.