
Standards for Creating Bibliographic Databases in Indian Academic Libraries under INFLIBNET Umbrella

Rajesh Chandrakar
Prem Chand

K Prakash
T A V Murthy

Abstract

In the field of library and information science, importance of cataloguing, classification, abstracting and indexing standards play a major role to represent the appropriate information of any document for patron's use. In beginning, these standards were used to the documents in the print formats like shelf list, catalogue card etc. Gradually, after the invention of the computer technology, representation of information in print format has changed into the electronic format. There with the development of technology, due to flood of information and variations in the representation of the information, were forced the profession to develop different bibliographic standards; till date almost more than 30 bibliographic standards has emerged. Thus, every country is following their own bibliographic standards for representing the print documents into the electronic format. Same way, India is also working in creating the bibliographic databases of the resources available in different academic libraries of the country. INFLIBNET (Information and Library Network) Centre, UGC, Ahmedabad has embarked upon this issue and working in war footing level to achieve the goal. The authors describe the bibliographic standards being used for creating databases by Indian Academic libraries under INFLIBNET umbrella.

Keywords : Bibliographic Standards, Information Exchange Format, INFLIBNET

0. Introduction

The concept of "standard" in the field of library and information science is interpreted as codes of cataloguing rules, classification schemes, subject headings and various other documents contained recommendations for good library practice, all of which are subject to varying interpretation. But, for another professionals, standards are extremely precise technical specifications for communication between software and machines in an automated information-processing environment.[1] For example, Operating Systems, TCP/IP etc.

Standards are documented agreements containing technical specifications or other precise criteria to be used consistently as rules, guidelines, or definitions of characteristics, to ensure that materials, products, processes and services are fit for their purpose

Field of bibliographic standards are not limited to the rules only but spread up to the authority files such as person names, corporate names, subject headings, classification schemes, series names, institution names, software and standardization form and format etc.

In beginning, during the initiation of automation to library professionals, there was not much focus on bibliographic standards and were creating databases in any formats. That affected to the automation of their libraries, which forced them to think over the issue. Now, gradually, the libraries and information centres are realizing the importance of bibliographic standards. Centres like Information and Library Network (INFLIBNET), Ahmedabad is playing important role in bringing all libraries together to work on a single platform of standards with linking under one huge national network i.e. UGC-Infonet. INFLIBNET

Centre is an Inter-University Centre of University Grants Commission (UGC), working on Academic library system in India. In 1990, UGC launched it as a programme for five years on pilot basis. On June 1996, the programme was recognized as a full-fledged Centre. To know more about the Centre, kindly visit the website <http://www.inflibnet.ac.in>.

UGC-Infonet is a project of UGC being implemented by INFLIBNET Centre in collaboration with ERNET (Education and Research Network), New Delhi. It is an Intra-Network of Indian universities and colleges across the country based on the Open IP platform, employing state-of-the-art technologies like IP or TV, enabling online response to queries. The network facilitates Intranet and Internet access to its users on high data rates.[2]

1. Standards for Bibliographic Database

Well-known simple definition of database is a set of records and record is a set of fields, where, field is a combination of different pieces of information. Bibliographic database is a database of resources available at libraries in print format. With the revolution of information technology in the field of libraries, bibliographic catalogue has come up into the form of bibliographic database. For making compatible with these changes, there, many library institutions have come up with different standards and formats such as CCF, AUSMARC, CANMARC, UKMARC, USMARC, UNIMARC, MARC21, AACR2, ISBD, LCSH, MeSH etc. Some of the Internationally recognized standards and formats being used by many libraries of different countries for creating databases. As per the INFLIBNET's guidelines following standards are recommended for use:

1.1 Standardized Bibliographic Formats

During the inception of the Programme, The Task Group on "Union Catalogue, Data Capturing and Standardization" was constituted by UGC. The Task Group was assisted by a Sub-Committee on implementation of Common Communication Format (CCF) for data capturing and data exchange. Later on, the Sub-Committee formalized some fields based on CCF and named it as "Guidelines for Data Capturing: a user manual"[3].

During the initial stage of automation, Micro CDS/ISIS software (developed by Unesco) were suggested to use for creating bibliographic databases based on the format recommended by the Centre. Simultaneously, other DBMS such as DBase, FoxPro, FoxBase etc. were also being used by different libraries for creating the databases. In addition, the Centre had also the library house keeping software called "ILMS – Integrated Library Management System" developed on COBOL language on Unix platform, and was distributed freely to the member libraries of the Centre.

Later in February 2000, Centre developed another library house keeping software called "SOUL – Software for University Libraries" on Power Builder language on Windows platform by using the client-server architecture to keep in mind the shortcomings being faced in ILMS software. The software is developed based on the format formalized by Sub-Committee of the Task Force.

Later, early 2001, The Ministry of Human Resource Development (MHRD), Govt. of India visualizes building up a national database of library holdings in the country based on the MARC21 format. Further, after approval the committee directed INFLIBNET Centre to implement the MARC21 format in all libraries of the country. In this regards, the Centre formed the MARC21 Core Group during December 2001, which formalized the fields for implementing it with the SOUL (Software for University Libraries – to know more about the software kindly visit to web address <http://web.inflibnet.ac.in/info/souInfo.jsp>) library house keeping software.[4] The main reason for preferring MARC21 over CCF and other formats was that most of the countries in the world use MARC and it is comprehensive having number of fields. Also there is

constant development taking place in MARC21 and is updated regularly whereas CCF has become stagnant [5].

In the first meeting MARC21 Core Group were come out with the following results

- ✍ MARC21 Code list for field identification. Available on following website:
- ✍ http://web.inflibnet.ac.in/info/dwld/marc21_codelist.pdf
- ✍ Field list for Name and Subject Authority Database. Accessible through:
- ✍ http://web.inflibnet.ac.in/info/dwld/auth_tag_lst.pdf
- ✍ Fields for the Indian Manuscript Database. Available at:
- ✍ <http://web.inflibnet.ac.in/info/dwld/finalmanuscript.pdf>

2. Standard for Rendering Information / Bibliographic Descriptions

AACR2R, ISBD and CCC are standards for bibliographic descriptions. But, INFLIBNET Centre suggests their participating libraries to follow the Anglo American Cataloging Rules-II (AACRII) Revised version for rendering bibliographic information.

3. Standardized Subject Headings

Different subject headings list are available, developed by different institutions such as LCSH, Sear's List of Subject Headings, MeSH, Thesaurus etc. But INFLIBNET suggests to its member library to use LCSH (for generalized subject headings) and Engineering Index, MeSH and thesaurus (for specialized Subject Headings).

4. Standard for Classification Schemes

There are various classification schemes have been developed by different organizations for classifying the books on different subjects. Functionally, most of the libraries are using the Dewey Decimal Classification (DDC) for assigning the Classification Number to the documents. But, under INFLIBNET umbrella, all the libraries are free to use any classification scheme for classifying the document, as it is very difficult for any library to reclassify the entire collection based on any new scheme. As per the library practice any classification scheme can be used as MARC21 has capability to adopt any kind of classification schemes.

5. ISO-2709 Information Exchange Format

This standard specifies the requirements for generalized interchange format that will accommodate data describing all forms of material. It describes a generalized structure designed specifically for exchange of data between processing systems and not necessarily for use as a processing format within systems. The format may be used for the interchange of records using various communication media.[6] INFLIBNET Centre recommends to follow this Exchange Format for contributing records to the Centre. The exchange format has implemented too with the in-house developed software SOUL for exchanging records. Physical structure of the implemented information exchange format is given below:

5.1 Physical Structure

General Structure of the Record:

Leader	Directory	Data Fields	R/S
--------	-----------	-------------	-----

R/S =Record Separator

5.1.1 Leader

Each record begins with a leader of fixed length of 24 character positions and is the first instance of the ISO 2709 format of a record. It contains:

- ✍ **Record length (5 characters):**
The length of the record includes the leader, directory, data fields and record separator.
- ✍ **Record Status Code (1 character):**
Always '0' for output files, ignored on input files.
- ✍ **Implementation Codes (4 characters):**
These codes are defined for each specific implementation of the ISO 2709 format. Second character position represents Bibliographic level codes (i.e. s- serials, m- single volume monographs, c- multivolume monographs, a-component part) and rest of the character positioned not used. Here, Bibliographic level is defined here as per the INFLIBNET's use. For the MARC21 based records, it follows the description defined for bibliographic level in MARC21 bibliographic format.
- ✍ **Indicator length (1 character):**
This is always '0', because here at INFLIBNET database indicator is not being used. But the value is '2', when we export records from SOUL to MARC21 in ISO 2709 format.
- ✍ **Subfield identifier length (1 character):**
'2' The subfield identifier length. This shows the number of characters being used for the subfield identifier.
- ✍ **Base address of data (5 characters):**
A decimal number indicating the number of control characters (length of leader plus directory and field separator) which precede the variable data fields. This number, when added to the starting address of the record, gives the address of the first character of the variable fields.
- ✍ **For user systems (3 characters):**
These character positions are not used. Always '000' in output files.
- ✍ **Length of "Length of Datafield" field (1 character):**
'4' the length of "Length of Datafield" in the directory. (Use of 4 characters permits datafields as long as 9,999 characters)
- ✍ **Length of "Starting character" position (1 character):**
'5' the length is "Starting character position" in the directory.
- ✍ **Rest (2 characters):**
These character positions are not used. Always sets on '00'.

5.1.2 Directory

A directory entry in this bibliographic record is made up of a tag, length-of-field, and field starting position. It is of variable length and consists of a series of fixed fields, referred to as "entries". One directory entry

is associated with each datafield present in the record. Each directory entry is 12 characters in length and is divided into three parts:

- ✍ Tag
- ✍ Length of datafield
- ✍ Starting character position

Tag: A three numeric ASCII character code identifying the datafield, which corresponds to the directory entry. In software, names of the fields are being used in alphabetical name, but during the export in ISO 2709 format and display on screen, it converts in three-numeric digit ASCII character code.

Length of datafield: A four digit number showing how many characters are occupied by the datafield, including datafield separator but excluding the record separator code if the datafield is the last field in the record.

Starting character position: A five-digit number giving the position of the first character of the datafield relative to the base address of data, i.e. the first character of the first of the datafields.

STRUCTURE OF A DIRECTORY ENTRY

TAG	FIELD LENGTH	STARTING CHARACTER POSITION
-----	--------------	-----------------------------

5.1.3 Datafields

A datafield consists of subfield and datafield separator:

- ✍ **Subfield :** A subfield consists of a subfield identifier followed by a data string, which is terminated by either another subfield identifier or a field separator. A subfield identifier consists of a subfield identifier flag i.e. cap (^) sign followed by one another character, that can be alphabet a to z (small or capital) and numeric 1 to 9, INFLIBNET Centre uses here small alphabets. For MARC21 based records "\$" graphical sign is used as a subfield identifier.
- ✍ **Datafield Separator :** The datafield separator i.e. hash (#) sign constitutes the final character of every datafield. ASCII value 30 (character 1/14 of ISO 646) is being used as a field separator in MARC21 based data.

A datafield, which has a single subfield, will be organised as follows:

SUBFIELD IDENTIFIER	SUBFIELD VARIABLE	FIELDSEPARATOR
2 character		1 characters

A datafield, which has two subfields, will be organised as follows:

FIRST SUBFIELD IDENTIFIER	FIRST SUBFIELD VARIABLE	SECOND SUBFIELD IDENTIFIER	SECOND SUBFIELD VARIABLE	FIELD SEPERATOR
2 characters		2 characters		1 character

5.1.4 Record Terminator / Separator

The record terminator is the final character of the record i.e. hash (#) sign. It follows the field separator of the final datafield of the record. ASCII 29 (character 1/13 of ISO 646) is being used as record separator in MARC21 based data.

6. Other Standards related to bibliographic database:

There are some other standards related to the bibliographic databases, are given below:

Z39.50 Retrieval protocol

Z39.50 refers to the International Standard, ISO 23950: "Information Retrieval (Z39.50): Application Service Definition and Protocol Specification", and to ANSI/NISO Z39.50. The standard specifies a client/server-based protocol for searching and retrieving information from remote databases. More can be found on <http://lcweb.loc.gov/z3950/agency/>.

OAI (Open Archive Initiative)

The Open Archive Initiative (OAI) for metadata harvesting is a new protocol dedicated to solving problems of digital library interoperability by defining simple protocols, most recently for the exchange of metadata. Although, Z39.50 Information Retrieval Protocol is also one of the harvesting protocols. Some digital library systems are using for retrieving and searching the records for example Maxwell System. It is crucial to explain, discuss, and disambiguate the concepts and terminology used among OAI implementers to harvesting approach to interoperability.

The harvesting protocol defined by the OAI is a request / response protocol with 6 request types, viz. *Identity*, *ListMetadataFormats*, *ListSets*, *GetRecords* *ListRecord*, and *ListIdentifiers*. It has defined in such a way to support a common set of principles and a technical framework to achieve interoperability.

Although, both the standards are not being used with the INFLIBNET database right now, but being explored to implement it in near future.

7. Purpose of the Bibliographic Standards:

Mostly, purpose of using these standards for creating bibliographic databases or suggesting using it at libraries of India, has following reasons behind it:

- ✍ Pursuing Universal Bibliographic Control
- ✍ Exchange of information among union databases for the purpose of global resource sharing
- ✍ Cooperation at the local level, transferring data from one system to another

8. Importance of the Bibliographic Standards

Following bibliographic standards for creating the databases produces following importance:

- ✍ It makes uniformity in records
- ✍ It provides better resource sharing between different resource centres
- ✍ It helps in exchanging of records and facilitates it without data loss
- ✍ It helps in creating national / International bibliographic union database
- ✍ It helps to adopt any system and provides platform interoperability

9. Conclusion

For libraries in India, it is very difficult to strict with any standard, due to libraries are not well recognized by their institutions and having lack of skilled manpower. A financial crunch with lack of skilled manpower is threat for the Indian libraries for keeping themselves with the pace of latest technology including the intruders like computer professionals. But to survive in the field, one has to go through the standards and control the quality in automation. Here INFLIBNET Centre is working as catalyst for the academic community helping into every area of their development, as a effect of that libraries are gradually adopting the new technology and keeping pace with the latest trends in the area. Further, through the human resource development activity in the Centre unskilled manpower is getting skilled with the help of different training on different specialized areas. Most importantly, recently, Centre has joined the National Information Standards Organization (NISO, USA –<http://www.niso.org>), which is a not-for-profit association accredited by ANSI, identifies, develops, maintains, and publishes technical standards to manage information in our changing and ever-more digital environment. This will help to keep us update in the standards areas, where INFLIBNET will also play major role in the development of global standards and influence it on Indian point of view. INFLIBNET is also representative of Bureau of Indian Standards (BIS) Technical Committee MSD5. This will help the Centre to educate the nation in the area of development of standards and its implementation into country.

10. References

1. Chiruppa, Siriporn. Standards for bibliographic control in Thailand. *In* 65th IFLA Council and General Conference held at Bangkok, Thailand on August 20 – August 28, 1999. <http://www.ifla.org/IV/ifla65/papers/139-155e.htm>
2. UGC-Infonet. <http://web.inflibnet.ac.in/info/ugcinfonet/ugcinfonet.jsp>
3. INFLIBNET Centre. Guidelines for Data Capturing: a user manual. Ahmedabad: INFLIBNET Centre, [1993]. 101p.
4. MARC Code List for Field Identification http://web.inflibnet.ac.in/info/dwld/marc21_codelist.pdf
5. INFLIBNET Newsletter Vol. 7, no. 1, 2001(March 2001) http://web.inflibnet.ac.in/res/news1_7_1.jsp#8
6. ANSI/NISO Z39.2 – 1994. Information Interchange Format. Bethesda: NISO Press, 1994. 12p.
7. Information and Library Network (INFLIBNET) Centre. <http://www.inflibnet.ac.in>
8. National Information Standards Organization (NISO) <http://www.niso.org>

About Authors



Rajesh Chandrakar is working as Scientific and Technical Officer at INFLIBNET, for last seven and half years. He is working with Database Development & Management Group, and Networking and Quality Control Group. He holds Bachelor degree in Science from Govt. Model College of Science, Raipur, Chhattisgarh and Master degree in Library and Information Science from Pt. Ravishankar Shukla University, Raipur. He also holds the Postgraduate Diploma in Computer Application from Pt. Ravishankar Shukla University, Raipur. He is Convenor of MARC21 Core Group of INFLIBNET Centre formed to implement the MARC21 in Indian academic universities. He is an alternate representative of Bureau of Indian Standards (BIS), Technical Committee MSD 5 and National Information Standards Organization (NISO, USA). He has more than 20 articles in his credit published at International and national level. Couple of articles has been cited in Application of New Technologies in Libraries, weekly citation. He is acting as Joint Convener of International CALIBER 2005 to be held at Cochin University of Science and Technology, Kochi.



Mr. K Prakash is working as Scientific/Technical Officer-I with INFLIBNET Centre since 1995. He has his basic degree in Science and Masters Degree in Library and Information Science from Karnatak University, Dharwad. He has qualified SLET. Pursuing research in Library Automation. He has done specialization course in "Information Technology Applications to Library and Information Services" from NCSI, IISc Bangalore. Before joining to INFLIBNET, he has worked in academic and industrial libraries. Presently he is working in Serials Union Database Development & Managing, and in addition to this he is involved in Training and other Activities of the centre. He has contributed several papers in Seminars and conferences. He is a life member of several professional bodies and he is managing digilib_India forum also. His areas of interests are Library Automation, Database Management, Information Retrieval, Organisation of e-resources, Digital Libraries and Training etc.



Mr. Prem Chand is Scientist–C and Group leader of Database Development and Management Group (DDMG) at INFLIBNET Centre. He holds Master degree in Political Science and Master in Library and Information Science and currently pursuing Ph D. Before joining INFLIBNET, he worked as Sr. Library and Information Assistant at Lal Bahadur Shastri National Academy of Administration (LBSNAA) Mussoorie. He has over 13 years of experience in Library profession. He has published several papers in various professional journals and conferences and edited CALIBER – 2003 conference proceeding. He has completed two major projects funded by NISSAT and WHO besides coordinated number of training programmes at Centre and at various universities. He is involved in development of union catalogue of academic libraries in India. His areas of interest are Bibliographic formats and standards, database management library automation, consortia, open Access and digital Library.



Dr. T.A.V. Murthy is Director of INFLIBNET and President of SIS. He holds B Sc, M L I Sc, M S L S (USA) and Ph.D. He carries with him a rich experience and expertise of having worked in managerial level at a number of libraries in many prestigious institutions in India including National Library, IGNC, IARI, Univ of Hyderabad, ASC, CIEFL etc and Catholic Univ and Casewestern Reserve Univ in USA. His highly noticeable contributions include KALANIDHI at IGNC, Digital Laboratory at CIEFL etc. He has been associated with number of universities in the country and has guided number of Ph.Ds and actively associated with the national and international professional associations, expert committees and has published good number of research papers. He visited several countries and organized several national and international conferences and programmes. E-mail : tav@inflibnet.ac.in