

Subject: Library and Information Science

Production of Courseware
 -Content for Post Graduate Courses



Paper No: 05 ICT for Libraries

Module : 05 Understanding Modern Library Standards: MARC 21 and Dublin Core



Development Team

Principal Investigator
&
Subject Coordinator

Dr. Jagdish Arora, Director
INFLIBNET Centre, Gandhinagar

Paper Coordinator

Dr Usha Mujoo Munshi, Librarian,
Indian Institute of Public Administration

Content Writer

Shri Dinesh Pradhan
Shri Miteshkumar Pandya

Content Reviewer

Dr Usha Mujoo Munshi, Librarian,
Indian Institute of Public Administration

Description of Module	
Subject Name	Library and Information Science
Paper Name	05 ICT for Libraries
Module Name/Title	Understanding Modern Library Standards: MARC 21 and Dublin Core
Module Id	LIS/ICTL-C/05
Pre-requisites	
Objectives	
Keywords	MARC 21, Dublin Core, Library Standards

I. Objectives

The module intends to impart knowledge of the following aspects of MARC21 and Dublin Core metadata standards:

- i) To understand basics of bibliographic standard;
- ii) Understand concept of MARC 21;
- iii) Basics of MARC tagging system;
- iv) Understand concept of Dublin Core;
- v) Basics of Dublin Core metadata elements;
- vi) To understand record structure of MARC record.

II. Learning Outcome

At the end of this module, participants will have understanding of MARC bibliographic standard and MARC tags and its use in integrated library management software. Participants will get basic idea of Dublin Core metadata standards used for preservation of digital contents. Participants will also be able to get a fair idea about the encoding of catalogue records using MARC21.

III. Structure of the Module

- 1 Introduction
- 2 Structure of MARC21 Record
- 3 Types of MARC21 Standard
- 4 Dublin Core
- 5 Summary
- 6 References

1. Introduction

The core of the automated system is the bibliographic information/metadata that describes the materials in the collection of libraries and provides search terms so that these materials can be discovered. It not only provides a means for the users to search the library collection, but also helps the library professionals in managing the library collections. Over the years, cataloguers have defined various rules for standardization of the bibliographic information related to the library records. The most prominent and widely-used library cataloguing rules is AACR-II which is quite comprehensive, but are restricted to choice and rendering of bibliographic elements, they do not delve into encoding of bibliographic elements in an automated library system. The MARC21 (Machine Readable Catalogue) formats are standards for the representation and communication of bibliographic and related information in machine-readable form. The MARC code does not provide an instruction on description of library materials, but is a standard format for bibliographic information into a computerized

record. It is used to create catalogue records in a way so that computers can read and interpret. Simple book has two types of information i.e. authority and bibliographic. The information from a catalogue card cannot simply be typed into a computer to become part of an automated catalogue. The computer needs some program or code to interpret the information that is available in specific format on a catalogue card. MARC standard has ability to encode the bibliographic information in machine-readable format.

The MARC21 record contains various "signposts," before each piece of catalogue information so that it can be identified easily. It is an important bibliographic standard that ensures effective and efficient retrieval of data through Online Public Access Catalogue (OPAC). It offers variety of tags to store and describe related information that can be used for searching a catalogued item.

2. Structure of MACR21 Record

A MARC record consists of three elements, i.e. i) record structure; ii) content designation; and iii) content of the record.

- i) **Record Structure:** The marc standard presents a generalized structure for bibliographic, authority and holding data. However, it does not specify the content of a record. In simple words, the record structure provides information on tags, indicators and data elements that are used to create the record.
- ii) **Content Designation:** Predefined codes that are used as prefix to the data elements within a record to support their manipulation as defined in the MARC 21 formats. The content designation is designed to support the function of formatting the data for display such as printing of card catalogue, production of machine-readable catalogues, printing of bibliographies, etc. It is also used for information retrieval of the records. For example, content designation for author is 100 and for title it is 245.
- iii) **Content:** Content of most data elements is defined and rendered in accordance to standards such as Anglo American Cataloguing Rules.

The information in a MARC record is stored in character form and records are coded in extended ASCII format. MARC21 record contains Leader, Directory and Variable fields. The Leader is a 24 character fixed field that stores important information about the MARC record required for processing of the record through computer. Directory is an index to the location of variable fields (control and data) within a record and starts at position 24. It has an entry for each MARC field that contains 3 characters tag, field length in 4 characters and starting position for each field in 5th characters. The variable fields are classified into variable control

fields and variable data fields. The content in MARC record is divided into Leader, Directory and Variable Fields.

2.1 Leader

The leader is first field in any MARC record and has a fixed length of 24 characters (octets positions 0-23). Only graphical characters of ASCII are allowed in the Leader. The structure of the leader defined in the MARC21 is given below. The numbers indicate the character positions occupied by each part of the leader.

Structure of the Leader in MARC 21 Records

- **Record Length** (character positions 00-04) contains a five-character ASCII numeric string equal to the length of the entire record, including itself and the record terminator. The five-character numeric string is right justified and unused positions contain zeros (zero fill). The maximum length of a record is 99999 octets.
- **Record Status** (character position 05) contains an ASCII graphic character which indicates the relation of the record to a file (e.g., new, updated, etc.).
- **Type of Record** (character position 06) contains an ASCII graphic character which specifies the characteristics and defines the components of the record.
- **Implementation-defined** (character positions 07-08) ANSI Z39.2 and ISO 2709 reserve character positions 07-08 for definition by a particular implementation. The individual MARC 21 formats define these character positions, if needed. Positions may contain only ASCII graphic characters. Any position not defined contains a blank.
- **Bibliographic Level** (bibliographic record, character position 07) contains an ASCII graphic character which also provides information about the components and characteristics of the record.
- **Kind of Data** (community information record, character position 07) contains an ASCII graphic character which also provides information about the components and characteristics of the record.
- **Type of Control** (bibliographic record, character position 08) contains an ASCII graphic character which also provides information about the components and characteristics of the record.

- **Character Coding Scheme** (character position 09) contains a code that identifies the character coding scheme used in a record.
- **Indicator Count** (character position 10) contains one ASCII numeric character specifying the number of indicators occurring in each variable data field. In MARC 21 records, the indicator count is always 2.
- **Subfield Code Length** (character position 11) contains one ASCII numeric character specifying the sum of the lengths of the delimiter and the data element identifier used in the record. In MARC 21 records, the subfield code length is always 2. The ANSI Z39.2 and ISO 2709 name for this data element is **identifier length**.
- **Base Address of Data** (character positions 12-16) contains five ASCII numeric characters that specify the first character position of the first variable field in the record. It is equal to the sum of the lengths of the leader and the directory, including the field terminator at the end of the directory. The number is right justified and unused positions contain zeros (zero fill).
- **Implementation-defined** (character positions 17-19) ANSI Z39.2 and ISO 2709 reserve character positions 17-19 for definition by a particular implementation. The individual MARC 21 formats define these character positions, if needed. Positions may contain only ASCII graphic characters. Any position not defined contains a blank.
- **Entry Map** (character positions 20-23) contains four single digit ASCII numeric characters that specify the structure of the entries in the directory.
- **Length of field** (character position 20) specifies the length of that part of each directory entry; in MARC 21 records, it is always set to 4.
- **Length of Starting-character Position** (character position 21) specifies the length of that part of each directory entry; in MARC 21 records, it is always set to 5.
- **Length of Implementation-Defined** (character position 22) specifies that part of each directory entry; in MARC 21 records, a directory entry does not contain an implementation-defined portion, and therefore, this position is always set to 0.
- **Undefined** (character position 23): This character position is undefined; it is always set to 0.

2.2 Directory

The Directory is an index to the location of data within a record, consisting of a series of fixed-length entries, with one entry for each variable field (control or data) presented in a record. Each Directory entry has 12 characters in length, starting with position zero (0) and ending with position 11. Each entry contains three portions i.e. field tag, field length and starting character position. Directory is located in the record at 24th position and it follows the leader. Directory for all types of MARC record is mandatory and it does not have any indicators and subfield codes.

Positions	Description
00-02	Tag
03-06	Field length
07-11	Starting character position

Directory entries for variable control fields appear first, sequenced by tag in increasing numerical order. The stored sequence of the variable data fields in a record does not necessarily correspond to the order of the Directory entries. Duplicate tags are distinguished only by the location of their respective fields within the record. The Directory ends with a field terminator character (ASCII 1E hex).

2.3 Variable Fields

2.3.1 Control Fields

The 00X fields in the MARC 21 formats are variable control fields. Variable control fields consist of data and a field terminator. They contain neither indicators nor subfield codes. Variable control fields contain either a single data element or a series of fixed-length data elements identified by relative character position.

2.3.2 Variable Data Fields

All fields except 00X are variable data fields. Four levels of content designation are provided for variable data fields in ANSI Z39.2:

- A three-character tag, stored in the directory entry.
- Indicators stored at the beginning of each variable data field, the number of indicators being reflected in Leader/10 (Indicator count).
- Subfield codes preceding each data element, the length of the code being reflected in Leader/11 (Subfield code count).
- A field terminator following the last data element in the field.

In variable data fields, the Indicators contain values conveying information that interprets or supplements the data found in the field. The MARC 21 formats specify two indicator positions at the beginning of each variable data field. Indicators are defined independently for each field. Parallel meanings are preserved, whenever possible. Indicator values are interpreted independently; meaning is not ascribed to the two indicators taken together.

Indicators may be any lowercase alphabetic or numeric character or a blank (#). Numeric values are defined first. A blank (#) is used in an undefined indicator position or to mean information not provided in a defined indicator position. The blank may have specific meaning when necessary for upward compatibility. The value 9XX is reserved for local implementation.

The variable field in MARC record may have 1 to 9,999 characters. Variable field is bifurcated in tags, indicators and subfields. Each variable field has three parts:

- Tags (3 digits numeric code)
- Indicators (maximum 2 digits)
- Subfield Codes (maximum 1 digit)

Tags - The 3 digit numeric tags are assigned to the fields that they represent. It is possible that the tag may appear more than once in a record, thus, it is labeled repeatable (R). The fields that are not appearing more than once are labeled as non-repeatable (NR). For example, a catalogue record can have several subject headings, so the tags for subject

headings added entries (6XX) are labeled repeatable (R) and title statement is labeled as non-repeatable (NR).

Main Entry Begins with 1 Subject Heading Begins with 6 Series Added Entry Begins with 8

Series Begins with 4 Added Entry Begins with 7

	Main Entries 1XX	Series 4XX	Subject Headings 6XX	Added Entries 7XX	Series Added Entries 8XX
Personal Name X00	Main Entry Personal Name 100		Personal Subject Heading 600	Personal Name - Added Entry 700	Personal Series- Added Entry 800
Corporate Name X10	Main Entry Corporate Name 110		Corporate Subject Heading 610	Corporate Name - Added Entry 710	Corporate Series - Added Entry 810
Conference Name X11	Main Entry Meeting Name 111		Conference Subject Heading 611	Meeting Name - Added Entry 711	Meeting Conference - Added Entry 811
Uniform Title X30	Main Entry Uniform Title 130		Uniform Title Subject Heading 630	Uniform Title - Added Entry 730	Uniform Title Series - Added Entry 830
Title X40		Series Title Entry 4XX			

Uniform Title ends with 30
Conference Name ends with 11
Corporate Name ends with 10
Personal Name ends with 00

Figure - 1 MARC21 Tags

Indicators - There are total 2 indicators for each field. Each indicator has a single character and positions of both the indicators follow each tag. Indicators are not applicable to first nine tags i.e. 001 to 009. One or both of these character positions may be used for indicators. In some fields, only the first or second position is used; in some fields, both are used; and in some fields, like Fields 020 and 300, it is never used. When an indicator position is not used,

that indicator is referred to as "undefined" and the position is left blank. It is the convention to represent a blank, or undefined, indicator position by the character "#".

Each indicator value is a number from 0 to 9. However, two indicators together may look like a 2-digit number; they really are two single-digit numbers. The allowable indicator values and their meanings are spelled out in the MARC 21 documentation. The following example shows, the first 3 digits are the tag number for title field (245 defines this as a title field) and extent to the next 2 digits i.e. 1 and 4 are indicator values. The 1 is the first indicator; 4 is the second indicator.

245 14 \$a Introduction to Digital Library /

\$c derived from epopathshala.

Subfield codes - All the data in each field (beginning with the 010 field) is divided into subfields, each of which is preceded by a delimiter-subfield code combination. The most common subfield codes used with each tag are shown. Each subfield code is preceded by the character \$, signifying as a delimiter.

Sample MARC record

```
01554cam a2200325 a
450000100090000000500170000900800410002690600450006792500440011295502600
015601000170041602000340043304000180046705000200048508200170050510000310
052224500800055325001130063326000640074630000190081050501040082965000350
093370000220096870000280099070000190101870000240103770000200106195200990
1081852004801180-13182198-20110930155247.0-030501s2003 ohu 001 0 eng -
a7bcbccorignewd1eocipf20gy-gencatlg-0 acquireb2 shelf copiesxpolicy default- apc14
2003-05-01 HANDCARRY to HLCDclh03 2003-05-01dlh39 2003-05-01elh44 2003-05-01
```

to Deweyaaa20 2003-05-02apx11 2004-02-20 bk rec'd, to CIP ver.flh45 2004-02-26glh45
 2004-02-26 to BCCDapv01 2004-06-02 copy 2 set to BCCDeCatRefProj he09 2010-09-24-
 a 2003050872- a0910608709 (set : alk. paper)- aDLCcDLCdDLC-00aZ696b.D52
 2003-00a025.4/31221-1 aDewey, Melvil,d1851-1931.-10aDewey decimal classification and
 relative index /cdevised by Melvil Dewey.- aEd. 22 /bedited by Joan S. Mitchell, Julianne
 Beall, Giles Martin, Winton E. Matthews, Jr., Gregory R. New.- aDublin, Ohio :bOCLC
 Online Computer Library Center,c2003.- a4 v. ;c25 cm.-0 av. 1. Manual. Tables -- v. 2.
 Schedules 000-599 -- v. 3. Schedules 600-999 -- v. 4. Relative index.- 0aClassification,
 Dewey decimal.-1 aMitchell, Joan S.-1 aBeall, Julianne,d1946--1 aMartin, Giles.-1 -
 aMatthews, Winton E.-1 aNew, Gregory R.- aTranscription and added entries for all
 editors per Desc Pol, ta03 9-1-89 (copied from ed. 21)-0 br-MRRhZ696i.D52 2003tCopy 2-
 mAlczv. 1-4-

3. Types of MARC21 Standard

MARC21 is grouped into five major categories based on data type i.e. bibliographic, authority, holding, classification, community. It is necessary to understand bibliographic, authority and holding tags related information as in day-to-day work all these tags are used.

3.1 Authority

MARC 21 fields for authority data is designed for storing information concerning the authorized forms of personal names, subject keywords, and subject subdivisions. The information which is repeatable and needs to be entered in the specific form without any spelling mistake is stored in the authority files.

The term personal name refers to:

Personal names (X00), corporate names (X10), meeting names (X11), names of jurisdictions (X51), uniform titles (X30), name/title combinations

The term subject refers to:

topical terms (X50) topical terms, geographic names, and

geographic names (X51) genre/form terms with subject subdivisions

genre/form terms (X55) chronological terms (X48)

names with subject subdivisions uniform titles with subject subdivisions (X30)

3.2 Bibliographic

The MARC21 fields for bibliographic data is designed to store bibliographic information about printed materials, computer files, maps, music, serials. Bibliographic data commonly includes titles, names, subjects, notes, imprint information, and information about the physical description of an item. The bibliographic format contains data elements for various types of materials i.e. Books, Continuing resources/Serials Publication, Computer files, Maps, Music, Visual materials, Mixed materials, etc. The bibliographic data refers to the following tags:

0XX: Numbers and Codes

1XX: Main Entry Fields

20X-24X: Title and Title-Related Fields

25X-29X: Edition, Imprint, Etc. Fields

3XX: Physical Description, Etc. Fields

4XX: Series Statement Fields

5XX: Note Fields

6XX: Subject Access Fields

700-75X: Added Entry Fields

76X-79X: Linking Entry Fields

800-840: Series Added Entry Fields

9XX: Locally-Defined Fields

3.3 Holding

The MARC21 format for holdings data is designed to store holdings information for three types of bibliographic items that are identified by a code available in Leader/06 (Type of record):

- **Single-part item** - An item which is complete in a single physical part, for example, a single map, a one-volume book, a computer file (Leader/06, code x)
- **Multipart item** - An item which is complete, or intended to be complete, in a finite number of separate physical parts, for example, a set of maps, a musical score and parts, a ten-volume encyclopedia, a multimedia kit, a manuscript collection (Leader/06, code v)
- **Serial item** - An item which is issued in successive parts at regular or irregular intervals and is intended to be continued indefinitely, for example, a journal (Leader/06, code y)

Following tags are used to store holdings information of various types of metadata Fields

- Captions and Pattern (fields 853-855)
- Enumeration and Chronology (863-865)
- Textual Holdings (866-868)
- Item Information (876-878)

4. Dublin Core Metadata Initiative

Dublin Core is a metadata standard used for describing digital documents. The standard named as +Dublin Core+, because it was evolved in a invitational workshop held at Dublin, Ohio during 1995 and core is added because all the elements defined in the standard are broad, generic and usable in describing wide range of digital resources. It is also known as standard for cross-reference description of digital resources. The Dublin Core metadata is a list of around 18 elements out of which 15 elements are used frequently. The elements are grouped together based on its characteristics into three major categories i.e. content, intellectual property and instantiation as mentioned in the figure -2 below.

Content	Intellectual Property	Instantiation
<ul style="list-style-type: none"> • Title • Subject • Coverage • Description • Type • Audience • Relation • Source 	<ul style="list-style-type: none"> • Creator • Publisher • Rights • Contributor 	<ul style="list-style-type: none"> • Date • Identifier • Format • Language

Figure -2 Dublin Core Elements

Each of these element stores specific pieces of information related to a digital object. A unique label is defined for each element which is syntactic specification of elements doe encoding schemas.

4.1 Title

Label: Title

Description of element: This element allows users to store title of a digital document. The title is a name given to the resource or the name of document by which it is known.

For example:

Title="Introduction to Web 2.0"

Title="Evaluation of Digital Library"

4.2 Subject

Label: Subject and Keywords

Description of element: This element allows users to insert and store subject keywords on the topic of the content of digital resource. Generally, subject or keyword describes the topic of the resource which is usually derived from the controlled vocabulary.

For example:

Subject="Semantic Web"

4.3 Coverage

Label: Coverage

Description of element: This element allows users to store extent or scope of the content of resource. Generally, it includes location or period i.e. a place name or geographic locations, period or date.

For example:

Coverage="2014-2015"

Coverage="Ahmedabad, Gujarat, India"

Coverage="21st century"

4.3. Description

Label: Description

Description of element: This element will allow users to store description about the topic elaborated in the digital resource. Description may include but is not limited to an abstract, table of contents, reference to a graphical representation of content or a free-text.

For Example:

Description="The digital object is created to create awareness among library community."

4.4. Type

Label: Resource Type

Description of element: This element allows users to store type of in which digital object is created.

To describe the physical or digital manifestation of the resource, use the FORMAT element.

For example:

Type="Video"

Type="Text"

Type="Image"

4.5 Audience

Label: Audience

Description of element: This element allows user to store class of an entity for whom the resource is intended or useful. A class of entity may be determined by the creator or the publisher or by a third party.

For example:

Audience="elementary school students"

Audience="ESL teachers"

Audience="deaf adults"

Audience="dumb children"

4.6 Relation

Label: Relation

Description of element: This element allows users to store a reference to a related resource.

Example:

Title="2 States: A Film by Abhishek Verman"

Relation=" 2 States: The Story of My Marriage" [Resource is a novel, written by famous Indian novelist Chetan Bhagat]

There are many other relationships which may be read from the documentation available at the official website of Dublin Core. These relationship includes **IsPartOf**, **IsVersionOf**, **IsFormatOf**, **HasFormat**, **IsReferencedBy**, **References**, **IsBasisFor**, **IsBasedOn**, **Requires**, etc.

4.7 Source

Label: Source

Description of element: This element allows users to store information about resource from which present resource is derived. The present resource may be derived from the Source resource in whole or part.

For example:

Source="330.54/MAN" [where "330.54/MAN" is the call number of the print version of the resource, from which the present version is created]

Source: Page no. 41-50 are digitized

4.8 Creator

Label: Creator

Description of element: This element allows users to store name of an entity primarily responsible for making the content of the resource. It could be a person, an organization, or a service.

For example:

Creator="Arora, Jagdish "

Creator="Manjunath, G.K"

Creator="INFLIBNET Centre"

4.9 Publisher

Label: Publisher

Description of element: This element allows users to store the name of entity responsible for making the resource available. It could be publisher, a person, an organization, or a service.

For example:

Publisher="Delhi University"

Publisher="Taylor & Francis"

4.10 Rights

Label: Rights Management

Description of element: This element allows users to store the information about rights held in and over the resource. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, etc.

For examples:

Rights="Access limited to members"

Rights="http://cs-tr.cs.cornell.edu/Dienst/Repository/2.0/Terms"

4.11 Contributor

Label: Contributor

Description of element: This element allows users to store the name of an entity responsible for making contributions to the content of the resource. It could be name of a person, an organization or a service.

4.12 Date

Label: Date

Description of element: This element allows users to store the information about the dates associated with an event in the life cycle of the resource.

For examples:

Date="1998-02-16"

Date="1998-02"

Date="1998"

4.13 Identifier

Label: Resource Identifier

Description of element: This element allows users to store reference to the resource which is available on specification location. This identifier could be Uniform Resource Identifier (URI), Uniform Resource Locator (URL), Digital Object Identifier (DOI) and International Standard Book Number (ISBN).

For example:

Identifier="http://www.niso.org/standards/resources/Metadata_Demystified.pdf/"

Identifier="ISBN 1-880124-59-9+

Identifier=" 10.1080/21670811.2014.1002513" [publisher number]

4.14 Format

Label: Format

Description of element: This element allows users to store information about physical or digital manifestation of the resource.

For example:

Format="image/gif"

Format="4 kB"

4.15 Language

Label: Language

Description of element: This element allows users to store language of the intellectual content of the resource.

For examples:

Language="en"

Language="en-US"

Sample record for Dublin Core

<dc:title>Handbook of DOI</dc:title>

<dc:creator>Manjunath, G K</dc:creator>

<dc:subject>Semantic Web</dc:subject>

<dc:description>This document is developed to create awareness about interactive and semantic web among library community in India.</dc:description>
<dc:publisher>Ess Ess Publication</dc:publisher>
<dc:contributor>Pandya, Miteshkumar</dc:contributor>
<dc:date>2014</dc:date>
<dc:type>Text</dc:type>
<dc:format>PDF</dc:format>
<dc:identifier> <http://dx.doi.org/10.1000/182> </dc:identifier>
<dc:source>Handbook of DOI</dc:source>
<dc:language>En</dc:language>
<dc:coverage>India</dc:coverage>

5. Summary

The module elaborates basic of MARC21 standard and its type such as authority, bibliographic and holding. It also describes records structure of MARC 21, i.e. leader, variable fields, variable data fields, etc. The module covers information of necessary tags which are required for day-to-day work. The module provides sample MARC record which helps to understand the structure of MARC record. The module also provides brief explanation of Dublin Core metadata standard and its elements for describing digital object.

References

1. Character Sets: General Character Set Issues: MARC 21 Specifications for Record Structure, Character Sets, and Exchange Media (Library of Congress). Retrieved September 29, 2015, from <http://www.loc.gov/marc/specifications/specchargeneral.html>
2. Dublin Core - The Elements. (2015). Retrieved November 30, 2015, from <http://dublincore.org/documents/2003/08/26/usageguide/elements.shtml>
3. Fritz, D. A., & Fritz, R. J. (2003). *MARC21 for everyone: a practical guide*. Chicago: American Library Association.
4. Leader (Record Structure). Retrieved September 29, 2015, from http://www.itsmarc.com/crs/mergedprojects/marcspec/marcspec/leader_record_structure.htm
5. The MARC 21 Formats: Background and Principles. Retrieved September 29, 2015, from <http://www.loc.gov/marc/96principl.html>

Excerpt taken from <http://www.loc.gov/marc> & <http://www.itsmarc.com> & <http://dublincore.org/documents/2003/08/26/usageguide/elements.shtml>