

curl

NEW CURL MARC 21 DATABASE:
Field Definitions

Version: 26/04/2010 @ 10:34

MIMAS
MANCHESTER INFORMATION & ASSOCIATED SERVICES

CONTENTS

**1. Database Control Fields.....	3
1.1 CURL Record Number (CRN : 1032).....	3
1.2 CURL Library Code (LIB : 56).....	3
1.3 Provenance (PROV : 6---).....	3
1.4 CURL Minimum Standard Flag (MSTD : 6---).....	4
1.5 Tags in Record (TAGS : 6---).....	4
1.6 Tag Count (TCNT : 6---).....	4
1.7 Subject Heading Count (SCNT : 6---).....	4
1.8 Subject Heading Flag (SFLG : 6---).....	4
1.9 Tag Total (TTOT : 6---).....	5
1.10 Date Record Loaded on CURL (DATE : 1011).....	5
1.11 Record Number on UKMARC Database (CRU : 6---).....	5
**2. Code Fields derived from Leader, 006, 007 and 008.....	6
2.1 Record Type (RTYP: 1001).....	6
2.2 Bibliographic Level (BIBL : 1021).....	6
2.3 Material Type (MTYP : 1031).....	6
2.4 Encoding Level (ENCL : 6---) <not yet complete>.....	7
2.5 Category of Material (CATM : 6---).....	7
2.6 Date Type/Publication Status (DTYP : 6---).....	8
Date Fields.....	8
2.7 Single Date (SDAT : 31, 3).....	9
2.8 Beginning Date (BDAT : 31, 4).....	9
2.9 Ending Date (EDAT : 31, 2).....	9
2.10 Country of publication (CPUB : 59).....	10
2.11 Government Publication Flag (GOVT : 6---).....	10
2.12 Conference Publication Flag (CONF : 6---).....	10
2.13 Language (LANG : 54).....	11
2.14 Microform Flag (MFRM : 6---) <not yet complete>.....	11
2.15 Electronic Form Flag (ELEC : 6---).....	11
**3. Control Number Fields.....	13
3.1 Original Local Record Number (CTRL : 6---).....	13
3.2 British National Bibliography Number (BNBN : 10).....	13
3.3 Library of Congress Control Number (LCCN : 9).....	13
3.4 National Bibliographic Agency Control Number (NCN : 6---).....	14
3.5 Other National Bibliography Number (NBN : 48).....	14
**4. Number and Code Fields.....	15
4.1 International Standard Book Number (ISBN : 7).....	15
4.2 International Standard Serial Number (ISSN : 8).....	15
4.3 Opus or Thematic Index Number (OPUS : 1030).....	15
4.4 Publisher's Number (PBNO : 51).....	16
4.5 Statement of Scale (Horizontal and Vertical) (SCAL : 1024).....	16
4.6 Standard Identifier (STID : 1007).....	17
**5. Classification Number and Related Fields.....	18
5.1 Dewey Decimal Classification Number (DDC : 13).....	18
5.2 Edition Number of DDC Schedules (EDDC : 6---).....	18
5.3 Library of Congress Classification Number (LCC : 16).....	18
**6. Basic Bibliographic Fields (Author, Title, Subject).....	19
Fields for Derived Searching.....	19

6.1.1	Author-Title Key (ATK : 1000, -, -, 3)	19
6.1.2	Title Key (TKEY : 4, -, -, 3)	20
6.2	Fields for Keyword Searching	20
6.2.1	Author (AUTH : 1003)	20
6.2.2	Uniform Title (UTI : 6)	20
6.2.3	Key Title (KEYT : 33)	21
6.2.4	Title (TITL : 4)	21
6.2.5	Series Title (SERT : 5)	21
6.2.6	Library of Congress Subject Heading (LCSH : 27)	21
6.2.7	Non-LC Subject Heading (SUBJ : 6--)	22
6.3	Fields for Browsing Indexes	22
6.3.1	Author Browse (ABRO : 1003)	22
6.3.2	Title Browse (TBRO : 4)	22
6.3.3	Subject Browse (SBRO : 21)	22
6.3.4	Name Browse (NBRO : 1002)	23
**7.	Fields in Alternate Scripts	24
7.1	Author in Alternate Script (XAU)	24
7.2	Title in Alternate Script (XTI)	24
7.3	Place of publication in Alternate Script (XPOP)	25
7.4	Name of Publisher in Alternate Script (XPUB)	25
7.5	Series Title in Alternate Script (XSER)	25
**8.	Other Fields	26
8.1	Place of publication (PPUB : 59)	26
8.2	Name of publisher (PUBR : 1018)	26
8.3	Pagination (PAGE : 6--)	26
8.4	Notes (NOTE : 63)	26
8.5	Uniform Resource Identifier (URI : 1209)	27
**9.	Display Fields	28
9.1	MARC Record (MREC)	28
9.2	Short Display (SDIS)	28
	Appendix I	29
	Appendix II	30
	Appendix III	31

INTRODUCTION

STATEMENT OF PURPOSE

The purpose of this document is to specify, by means of detailed field definitions, how we intend to store and index incoming MARC 21 bibliographic data in the new CURL database.

PRESENTATION

Generally, database field definitions are presented in this document as follows:

Database Field Name (Abbreviation : Z39.50 Use Attribute)

Reason for field (if not self-evident) and/or preliminary remarks.

Source fields: Specifies the MARC field(s) used to create the database field.
 Construction: Describes the manner in which the MARC field contents are manipulated in order to create/construct the database field.

PRE-PROCESSING

We are proposing to prepare incoming bibliographic records for loading into the database as follows:

- (1) We verify the data and generate report files.
- (2) We convert records from exchange format to MARC format.
- (3) We create an 000 local field with the following subfields:
 - \$a contains the Record Status
 - \$b contains the Record Type
 - \$c contains the Bibliographic Level
 - \$d contains the Date Record Loaded
 - \$e contains the CURL Minimum Standard Flag (optional)
 - \$f contains the Provenance code
- (4) We create an 001 containing the unique CURL record number.
- (5) We create an 003 containing the unique MARC 21 Organisation Code for CURL.
- (6) We create an 035 of the form "(CCCC)CTRL" where CTRL is the incoming record's Original Local Control Number and CCCC is the MARC 21 Organisation Code of the library sending us the record.
- (7) We create an 049 local field with the following subfields:
 - \$j contains "CU"
 - \$k contains the CURL Record Number
 - \$l contains the Provenance code
 - \$m (optional) contains 2 kinds of data:
 - (i) LC records: "+" means updated CIP record
 "-" means CIP record
 "%" means in process record
 - (ii) CURL records: CURL Minimum Standard Flag.

ON EXPORT

When an end-user downloads a record from the database:

- (1) We drop the 000 local field.
- (2) We create an 005 containing the date and time of the latest record transaction.
- (3) We convert the record from MARC format to exchange format. (Record includes 001, 003, 005, 035, 049 and all other MARC fields.)

INDEXED CHARACTERS

The characters which are indexed are a subset of the MARC 21 character set. The indexed characters are: A-Z 0-9 _ . @ - + & |

Thus, lowercase a-z are indexed as uppercase characters. Latin letters, which are based on the core set of 26 letters, are indexed as their nearest A-Z equivalent (e.g., Welsh lowercase “ŵ” {LATIN SMALL LETTER W WITH CIRCUMFLEX} is indexed as uppercase W.) Other characters, which are not on the list of indexed characters, are treated in most cases as if they were blank characters.

TYPOGRAPHICAL CONVENTIONS

The following typographical conventions are used in this document:

- # - The graphic symbol # is used for a blank character. (The blank cannot be indexed in the database but is used occasionally in MARC coded fields to convey meaning. If we want to search for blank (see the Encoding Level field, in 2.4, for example) then it needs to be treated in a special way.)
- | - The graphic symbol | is used for a fill character and indicates there has been no attempt to code a particular coded field.

~~MARC field~~ - The MARC field was considered for indexing but ultimately rejected.

OBSOLETE CONTENT DESIGNATORS

There are some references to obsolete content designators in this document. Obsolete content designators may appear in MARC records that were created prior to the time they were made obsolete. An obsolete content designator is not used in new records.

LOCAL CONTENT DESIGNATORS

There are some references to local content designators in this document. Local content designators may appear in records that originate from the U.S. or Canada prior to the creation of MARC 21. For more information, *See* MARC 21 Bibliographic Format, Appendix H.[‡]

ABBREVIATIONS

cp = character position
char = character

[‡] Unless otherwise stated, *See* references in this document refer to *MARC 21 Format for Bibliographic Data*, the 1999 print edition, including Updates 1-3 (2000-2002).

****1. Database Control Fields**

Fields for controlling basic database operation and for describing MARC tags in records.

1.1 CURL Record Number (CRN : 1032[†])

Unique identifier of record. The number will vary in length and is structured as follows:

XXOOO...O

where XX indicates the record source (see 1.2 below) and OOO...O indicates the Original Local Control Number (see 3.1).

E.g., 982001362615 would be the CRN for record 2001362615 from Library of Congress
17b10030933 would be the CRN for record b10030933 from Newcastle

Source field: 001
Construction: Take field contents as is.

1.2 CURL Library Code (LIB : 56)

Three-character code of the library contributing the record. See Appendix 1 for a list of codes.

Source field: 001
Construction: Convert first two digits of CRN to three-char code.

1.3 Provenance (PROV : 6---)

One-character code indicating the holder of the intellectual property right in the record. Values are:

Code	Definition
a	BLCMP
b	British Library
c	CURL
d	Other sources
l	Library of Congress
o	OCLC
r	RLG
s	SLS

Source field: 000 \$f
Construction: Take subfield contents as is.

[†] See Appendix II for more information about Z39.50 use attributes.

1.4 CURL Minimum Standard Flag (MSTD : 6---)

Flag which indicates the quality of a record. Values are: 12|.

“1” means record is substandard and is not displayed in search results.

“2” means record meets CURL minimum standard.

“|” means there has been no attempt to code the field.

```
Source field: 000 $e
Construction: LC records: if $e contains "+", set MSTD to 2
              CURL records: if $e contains "2", set MSTD to 2
                          if $e contains "1", set MSTD to 1
                          else set MSTD to |
```

1.5 Tags in Record (TAGS : 6---)

List of three-digit tags found in record.

Source fields: All MARC fields (including local fields).

Construction: Each tag needs listing only once, even if there is more than one in the record.

1.6 Tag Count (TCNT : 6---)

Count of different tags in record.

Source fields: All MARC fields (excluding local fields).

Construction: Count each different tag only once, even if there is more than one in the record.

1.7 Subject Heading Count (SCNT : 6---)

Count of 6XX tags in record.

Source fields: 6XX fields

Construction: Count each 6XX tag present in the record. Set SCNT to 0 if there are no 6XX tags.

1.8 Subject Heading Flag (SFLG : 6---)

Required to support limiting a search to records which contain 6XX fields.

Source fields: 6XX fields

Construction: If SCNT > 0 then set SFLG to "y".

1.9 Tag Total (TTOT : 6---)

Total number of tags in record.

Source fields: All MARC fields (excluding local fields).
Construction: Count all tags, except those in local fields.

1.10 Date Record Loaded on CURL (DATE : 1011)

Source fields: None
Construction: Generate system date stamp (YYYYMMDD).

1.11 Record Number on UKMARC Database (CRU : 6---)

Unique identifier of record on CURL UKMARC database. The number is fixed in length at 12 digits and is structured as follows:

XX1234567890

where XX indicates the record source and 1234567890 indicates the Original Local Control Number compressed to fit into 10 chars. The field is of interest to MIMAS only.

Source field: 001
Construction: Construct number from CRN.

****2. Code Fields derived from Leader, 006, 007 and 008**

2.1 Record Type (RTYP: 1001)

Required to support limiting a search by Record Type. Current MARC codes are: acdefgijkmoprt. Obsolete MARC codes are: bhn.

Source field: Leader cp6
 Construction: Take data as is, including the three obsolete codes.
 Convert # to |.[†]

2.2 Bibliographic Level (BIBL : 1021)

Required to support limiting a search by Bibliographic Level, and normally used in conjunction with a search also limited by Record Type. Only used in a few specialised cataloguing situations. Valid codes are: abcdims.

Source field: Leader cp7
 Construction: Take data as is, but convert # to |.

2.3 Material Type (MTYP : 1031)

Required to support limiting a search by Material Type. Values are: BK, CR[‡], CF, MP, MU, VM and MX (i.e., Books, Continuing resources, Computer files, Maps, Music, Visual materials and Mixed materials). Besides these 7 values, an 8th value (ZZ) is needed when the material type cannot be derived from the source fields alone. See: MARC 21 Concise Bibliographic: Introduction (<http://lcweb.loc.gov/marc/bibliographic/ecbdintr.html>). See also: MARC 21 Bibliographic Format, Introduction, p. 1-2.

Source fields: Leader cp 6 and 7
 006 cp 0
 Construction: There are 2 processes as follows:

Process no. 1

```
if Leader cp6 = a and Leader cp7 = a,c,d or m then set MTYP to BK
if Leader cp6 = a and Leader cp7 = b,i or s then set MTYP to CR
if Leader cp6 = a and Leader cp7 = any other value, then set MTYP to ZZ
```

```
if Leader cp6 = c then set MTYP to MU
                        d then set MTYP to MU
                        e then set MTYP to MP
                        f then set MTYP to MP
                        g then set MTYP to VM
                        i then set MTYP to MU
                        j then set MTYP to MU
                        k then set MTYP to VM
                        m then set MTYP to CF
                        o then set MTYP to VM
```

[†] We do not have to convert # to | but this may be better than leaving # as is.

[‡] Continuing resources (CR) is the new term for Serials (SE).

```

p then set MTYP to MX
r then set MTYP to VM
t then set MTYP to BK

```

For the obsolete codes:

```

if Leader cp6 = b then set MTYP to MX
    h then set MTYP to ZZ
    n then set MTYP to VM

```

Else set MTYP to ZZ.

Process no. 2

For each 006 present:

```

if 006 cp0 = a then set MTYP to BK
if 006 cp0 = s then set MTYP to CR

```

```

if 006 cp0 = c then set MTYP to MU
    d then set MTYP to MU
    e then set MTYP to MP
    f then set MTYP to MP
    g then set MTYP to VM
    i then set MTYP to MU
    j then set MTYP to MU
    k then set MTYP to VM
    m then set MTYP to CF
    o then set MTYP to VM
    p then set MTYP to MX
    r then set MTYP to VM
    t then set MTYP to BK

```

Else set MTYP to ZZ.

2.4 Encoding Level (ENCL : 6---) <not yet complete>

Required to support limiting a search by Encoding Level. Valid MARC 21 codes are: #1234578uz. OCLC has some additional codes (EJKLM) and we have not yet decided what to do with records using them.

Source field: Leader cp17
Construction: For contributing libraries which use the code, take data as is, except for records coded with #. This code indicates "Full level" and we are proposing to convert it to "f" because blank cannot be indexed.

2.5 Category of Material (CATM : 6---)

Required to support limiting a search by Category of Material. MARC 21 codes are: acdfghkmoqrstvz (i.e., Map, Computer file, Globe, Tactile material, Projected graphic, Microform, Nonprojected graphic, Motion picture, Kit, Notated music, Remote-sensing image, Sound recording, Text, Videorecording, Unspecified, respectively). A field of limited usefulness since few libraries create 007s.

Source fields: 007 cp0
 Construction: Take data as is. Convert # to |.

2.6 Date Type/Publication Status (DTYP : 6---)

Occasionally needed by library staff to limit searches. Valid codes are: bcdeikmnpqrstu|.

Source field: 008 cp6
 Construction: Take data as is, but convert # to |.

Date Fields

End-users need to be able to limit a search by a single date and by a range of dates. In order to support this the database requires three date fields: Single Date (SDAT), Beginning Date (BDAT) and Ending Date (EDAT). A record loaded on CURL will have either a Single Date or a Beginning Date/Ending Date pair, but not all three fields. *Exception:* When DTYP = m, we create all three date fields because this would facilitate searching for multipart items by a single date or by a range of dates.

At search time, a limit to a single date, X, would retrieve records with SDAT = X only. Whereas a limit at search time to a range of dates, A-Z, would retrieve records with ((A ≤ EDAT) AND (BDAT ≤ Z)) OR ((A ≤ SDAT) AND (SDAT ≤ Z)).

Construction: These three fields are constructed as follows:

Main Process

Let DATE1 = 008 cp7-10 and let DATE2 = 008 cp11-14.

If DTYP = b and DATE1 ≠ ##### and DATE2 = ##### then execute StoreDate1.
 If DTYP = b and DATE1 = ##### and DATE2 ≠ ##### then set BDAT to 0 and execute StoreEndDate.
 If DTYP = c or u then set EDAT to 9999 and execute StoreBeginDate.
 If DTYP = d, i, k, m or q then execute both StoreBeginDate and StoreEndDate.
 If DTYP = m then execute StoreDateTypeM.
 If DTYP = e, p, r, s or t then execute StoreDate1.
 Else omit all three date fields.

StoreDate1

If DATE1 contains u/U chars then copy DATE1 to DATE2 and execute both StoreBeginDate and StoreEndDate
 else set SDAT to DATE1.

StoreDateTypeM

Convert u/U chars in DATE1 to zeros.
 Set SDAT to DATE1.
 Set BDAT to DATE1.
 Execute StoreEndDate.

StoreBeginDate

Convert u/U chars in DATE1 to zeros.
Set BDAT to DATE1.

StoreEndDate

Convert u/U chars in DATE2 to nines.
Set EDAT to DATE2.

N.B.: Certain kinds of dates have not been indexed:

- (1) Date of production/recording session (i.e., DATE2 when DTYP = p)
- (2) Date of original publication (DATE2 when DTYP = r)
- (3) Copyright date (DATE2 when DTYP = t). Note, however, that a single copyright date recorded in DATE1 (which is common practice) would be indexed in SDAT as expected.

These dates could be indexed, possibly together in an Other Date field, if it was thought that this would be useful.

2.7 Single Date (SDAT : 31, 3)

This field is required to support limiting a search by date of publication, distribution, and so on. This field always contains a single, complete date, either known or probable. All 4 digits must be recorded in the source field; that is, missing digits, represented by u/U chars, are not permitted in SDAT. If the source field contains missing digits then SDAT is not created and a BDAT/EDAT pair is created instead. *Exception:* If the source field contains missing digits and DTYP = m, then convert missing digits to **zeros** and create SDAT and BDAT.

Source field: 008 cp7-10

2.8 Beginning Date (BDAT : 31, 4)

Required to support limiting a search by date of publication, distribution, and so on. This field always contains the beginning date of a range of dates. Missing digits, represented by u/U chars in the source field, are converted to **zeros**.

Source field: 008 cp7-10

2.9 Ending Date (EDAT : 31, 2)

Required to support limiting a search by date of publication, distribution, and so on. This field always contains the ending date of a range of dates. Missing digits are converted to **nines**.

Source field: 008 cp11-14 (normally)
008 cp7-10 (if cp11-14 is empty and cp7-10 contains u/U chars)

2.10 Country of publication (CPUB : 59)

Required to support limiting a search by code for country of publication.

Source fields: 008 cp 15-17
044 \$a

Construction: Take two- or three-char codes from fields, as is.

2.11 Government Publication Flag (GOVT : 6---)

Flag indicating that the item is a government publication.

Source fields: 008 cp 28, except for Mixed materials and Music
006 cp 11, except for Mixed materials and Music

Construction: There are 2 processes as follows:

Process no. 1

if Leader cp6 ≠ b, c, d, i, j or p
and 008 cp 28 ≠ 0, #, u or |
then put "y" in this field.

Process no. 2

For each 006 present:
if 006 cp0 ≠ c, d, i, j or p
and 006 cp 11 ≠ 0, #, u or |
then put "y" in this field.

2.12 Conference Publication Flag (CONF : 6---)

Flag indicating that the item is a conference publication.

Source fields: 008 cp 29, but only for Books and Continuing resources
006 cp 12, but only for Books and Continuing resources

Construction: There are 2 processes as follows:

Process no. 1

if Leader cp6 = a
and 008 cp 29 ≠ 0, # or |
then put "y" in this field.

Process no. 2

For each 006 present:
if 006 cp0 = a or s
and 006 cp 12 ≠ 0, # or |
then put "y" in this field.

2.13 Language (LANG : 54)

Required to support limiting a search by language of material. In newer records, multiple language codes are recorded in separate subfields, e.g., \$aeng\$arurus\$afre. In older records, multiple language codes were allowed to be recorded in a single subfield, e.g., \$aengrusfre.

Source fields: 008 cp 35-37
 041 \$a,b,d,e,f,g,h
 Construction: There are 2 processes as follows:

Old-Style Records

Check that the number of chars in each subfield is a multiple of 3, and then extract each code from the subfield. Take data as is, but convert ### to "+++". (Code ### means the item has no sung, spoken, or written textual content. See MARC 21 Bibliographic Format, 008-All materials, p. 12.)

New-Style Records

Extract each code from subfield. Take data as is, but convert ### to "+++".

2.14 Microform Flag (MFRM : 6---) <not yet complete>

Required to support limiting a search to microform publications. (Needs to be rewritten to take into account that 006 is repeatable.)

Source fields: 006 cp6, if MTYP = BK, MU, MX or CR
 006 cp12, if MTYP = MP or VM
 007 cp0
 008 cp23, if MTYP = BK, MU, MX or CR
 008 cp29, if MTYP = MP or VM
 245 \$h
 300 \$a
 Construction: If 006 cp6 = a, b or c AND MTYP = BK, MU, MX or CR
 or if 006 cp6 = a, b or c AND MTYP = MP or VM
 or if 007 cp0 = h
 or 008 cp23 = a, b or c AND MTYP = BK, MU, MX or CR
 or if 008 cp 29 = a, b or c AND MTYP = MP or VM
 or if 245 \$h = "[microform]"
 or if 300 \$a contains "aperture card" or "microfiche" or
 "microfilm" or "microopaque"
 then set MFRM to "y".

2.15 Electronic Form Flag (ELEC : 6---)

Required to support limiting a search to materials in electronic form.

Source fields: 008 cp 23/29, except for Computer files
 006 cp 6/12, except for Computer files
 Construction: There are 2 processes as follows:

Process no. 1

if Leader cp 6 = a,c,d,i,j,p or t AND 008 cp 23 = s
else if Leader cp6 = e,f,g,k,o or r AND 008 cp 29 = s
then put "y" in this field.

Process no. 2

For each 006 present:
if 006 cp0 = a,c,d,i,j,p,s or t AND 006 cp 6 = s
else if 006 cp0 = e,f,g,k,o or r AND 006 cp 12 = s
then put "y" in this field.

****3. Control Number Fields**

For these control number fields, each subfield specified below contains a number which needs to be indexed separately. This is different from the way we treat data in non-numeric fields.

3.1 Original Local Record Number (CTRL : 6---)

Source field: 035 \$a
 Construction: The form of the 035 field is "035##\$a(CCCC)CTRL" where CCCC is a MARC 21 Organisation Code and CTRL is the Original Local Record Number. Extract the number following the code.

3.2 British National Bibliography Number (BNBN : 10)

Each number exists in two formats, a BL format and an LC format. We have decided to index each BNB number twice, once in each format, in order to make searching more flexible.

Source field: 015 \$a
 Construction: Verify that number is a BNB number. Then construct two numbers as follows:

Numbers in BL format

Numbers are 8 chars long. A number begins with lowercase "b", followed by 2 chars for year of processing (the first of which may be alphabetic), followed by a 5-digit processing number (right justified with leading zeros). First char of 5-digit processing number is sometimes a letter.

E.g., Number in \$a is indexed as b9204551

E.g., Number in \$a is indexed as b97Y9761

E.g., Number in \$a is indexed as bA300001

Numbers in LC format

Numbers vary in length. A number begins with uppercase "GB", followed by 2 chars for year of processing, followed by a hyphen, followed by a processing number which consists of 1 to 5 digits (without leading zeros). As above, the first char of a 5-digit processing number is sometimes a letter.

E.g., Number in \$a is indexed as GB92-4551

E.g., Number in \$a is indexed as GB97-Y9761

E.g., Number in \$a is indexed as GBA3-1

3.3 Library of Congress Control Number (LCCN : 9)

Each number exists in two formats, a machine format (e.g., 76001332 or 2002087765) and a print format (e.g., 76-1332 or 2002-87765). Plus the structure of the number changed on 1 Jan. 2001, so there are old-style numbers and new-style ones. We have decided to index each LCCN twice, once in each format, in order to make searching more flexible.

N.B. The prefix lc/LC, currently used in 010 of CURL UKMARC records, will not be carried in 010 of CURL MARC 21 records. This is because it is not a valid Library of Congress prefix.

Source field: 010 \$a,z
 Construction: Right justify alphabetic prefixes (in order to facilitate keyword searching). Drop supplement numbers, suffixes and revision dates from old-style numbers. For each number in the source field, construct two numbers as follows:

Numbers in Machine Format

Numbers are fixed in length; old-style numbers are 11 chars and new-style ones are 12. Old-style numbers begin with a 3-char alphabetic prefix (ppp), followed by an 8-digit number. New-style numbers begin with a 2-char alphabetic prefix (pp), followed by a 10-digit number.

E.g., An old-style number is indexed as ppp76001332

A new-style number is indexed as pp2002087765

Numbers in Print Format

Numbers vary in length. Create numbers from numbers in machine format. Old-style numbers begin with a 3-char alphabetic prefix (ppp), followed by 2 digits for year of record creation, followed by a hyphen, followed by a serial number which consists of 1 to 6 digits (without leading zeros). Similarly, new-style numbers begin with a 2-char alphabetic prefix (pp), followed by 4 digits for year of record creation, followed by a hyphen, followed by a serial number which consists of 1 to 6 digits (without leading zeros).

E.g., An old-style number is indexed as ppp76-1332

A new-style number is indexed as pp2002-87765

3.4 National Bibliographic Agency Control Number (NCN : 6---)

Source field: 016 \$a,z
 Construction: Take data as is.

3.5 Other National Bibliography Number (NBN : 48)

Source field: 015 \$a
 Construction: If number is not a BNB number, take data as is.

****4. Number and Code Fields**

For these number and code fields, each subfield specified below contains a number which needs to be indexed separately. This is different from the way we treat data in non-numeric fields.

4.1 International Standard Book Number (ISBN : 7)

Source fields: 020 \$a, z
~~534 \$z~~
~~556 \$z~~
~~581 \$z~~

Construction: The number is normally carried in the MARC record without embedded hyphens. Extract number from first ten chars of subfield. Ignore all other chars.

4.2 International Standard Serial Number (ISSN : 8)

Source fields: 022 \$a, y, z
~~247 \$x~~
~~400 \$x [LOCAL to U.S.]~~
~~410 \$x [LOCAL to U.S.]~~
~~411 \$x [LOCAL to U.S.]~~
~~440 \$x~~
~~490 \$x~~
~~700 \$x~~
~~710 \$x~~
~~711 \$x~~
~~730 \$x~~

Construction: The number is normally carried in the MARC record with a hyphen between the 4th and 5th digits. Leave hyphen as is if present in source field or insert it if missing.
 E.g., 1234-5678 is indexed as is
 44445555 is indexed as 4444-5555

4.3 Opus or Thematic Index Number (OPUS : 1030)

Required to support searching for these numbers in music cataloguing records. Three types of numbers are carried in subfield \$n of music uniform titles: serial, opus and thematic index numbers. The problem is that serial and opus numbers may be coded together in the same subfield. We have decided to keep things simple by indexing the entire subfield. For record retrieval, we could facilitate searching by making the ADJ (adjacency) operator the default search operator for this field. (This will not always work but should work in the majority of cases.)

Source fields: 240 \$n, but only for Music
~~505~~

700 \$n, but only for Music

Construction: Take data from source fields only if MTYP = MU.

Normalise data as follows:

1. Replace each mark of punctuation with a #
2. Insert # between "no" and a "number"
3. Insert # between "op" and a "number"

4. Insert # between a thematic index abbreviation[†] and a "number"

A "number" may consist of an integer, an integer followed by a single letter, or (occasionally) a Roman numeral.

E.g., "no 1 op 2 no 1" is indexed as 6 words. A search on "op 2" would retrieve the record because the search term would be interpreted as "op ADJ 2". Note that a search on "op 1" would not retrieve the record.

E.g., "Hob XVI 27" is indexed as 3 words.

E.g., "D950" is indexed as 2 words ("D 950").

N.B.: Normalisation procedure will not always work.

4.4 Publisher's Number (PBNO : 51)

Required to support searching on Issue number, Matrix number, Plate number, Publisher's number, Other music number, Videorecording number and Other publisher's number.

Source fields: 028 \$a
 260 \$d [LOCAL to U.S.]
 262 \$k,1 [LOCAL to U.S.]
 305 \$m,n [OBSOLETE]
 Construction: Take data as is. (Data may include &, ., -, # and /)
 E.g., B. & H. 8797 is indexed as 4 words. ("&" and "." are indexed chars.)
 R.10150E. is indexed as 1 word.
 CV 70.089/03 is indexed as 3 words "CV 70.089 03".
 ("/" is not an indexed char.)

4.5 Statement of Scale (Horizontal and Vertical) (SCAL : 1024)

Required to support searching for these numbers in map cataloguing records.

Source fields: 034 \$b,c
 255
 Construction: Take data as is. E.g., 96000 is indexed as is.

[†] A list of "authorised" thematic index abbreviations is forthcoming and will be based on the list on p. 181-182 in: Smiraglia, Richard P. *Describing music materials: a manual for descriptive cataloging of printed and recorded music, music videos, and archival music collections*. 3rd ed., rev. and enl. Lake Crystal, Minn.: Soldier Creek Press, 1997.

4.6 Standard Identifier (STID : 1007)

Required to support searching for various standard identifiers, as specified in the Bath Profile, section 5.A.1.14.

Source fields: 020 \$a,z
022 \$a,y,z
024 \$a,z
027 \$a,z

Construction: Take data as is, but treat data in 020 and 022 as described above (in Sections 4.1 and 4.2). Special chars (those which are not on the list of indexed chars) are treated by the database software as if they were #.

E.g., 8756-2324(198603/04)65:2L.4:QTP:1-P is indexed as several words as follows:
8756-2324 198603 04 65 2L.4 QTP 1-P

****5. Classification Number and Related Fields**

5.1 Dewey Decimal Classification Number (DDC : 13)

Source field: 082 \$a
Construction: Remove all # and "/" chars. Leave "." as is.
E.g., 780/.9429/09034 is indexed as 780.942909034

5.2 Edition Number of DDC Schedules (EDDC : 6---)

Source field: 082 \$2
Construction: Take data as is.

5.3 Library of Congress Classification Number (LCC : 16)

Source fields: 050 \$a
055 \$a only if 2nd indicator = 0, 1, 3 or 4
090 \$a
Construction: Parse data and change "." to # only if it precedes a
letter.
E.g., HF5549.5.R44 is indexed as HF5549.5 R44
DC198.A1 is indexed as DC198 A1

****6. Basic Bibliographic Fields (Author, Title, Subject)**

Required to support basic bibliographic searching. These fields are basically textual in nature and, unless otherwise stated, they are constructed as follows:

Construction: Replace each subfield delimiter/code pair with a single #. For titles, remove non-filing chars as specified. (The code "X?" below means remove non-filing chars as specified by the **first** indicator. Similarly, the code "?X" means remove non-filing chars as specified by the **second** indicator.) Then take data as is.

Fields for Derived Searching

6.1.1 Author-Title Key (ATK : 1000, -, -, 3)

Required to support searching for Author-Title keys of the form (4-4). Author and title elements should be separated by a hyphen. E.g.,

shak-haml is the Author-Title key for Hamlet / by Shakespeare.

wu-six is the Author-Title key for Six contemporary dramatists / by Wu.

Source fields: 245 \$a (for title element)
 100 \$a (these, for author element)
 110 \$a
 111 \$a
 700 \$a
 710 \$a
 711 \$a

Construction: There are 2 processes as follows:

Author element

Use the first four alphanumeric chars (ignoring spaces and punctuation chars)--or as many chars as are available, if less than 4.

St. John -> stj

O'Brien -> obri

Title element

Remove non-filing chars (as specified by the second indicator) from the start of the title. Remove apostrophes and replace all non-alphanumeric chars by a single space (so hyphenated words become 2 words). Use the first word of the modified title (or as many chars as are available, if less than 4).

6.1.2 Title Key (TKEY : 4, -, -, 3)

Required to support searching for Title keys of the form (3-2-2-1). E.g.,

tal-of-tw-c is the Title key for A tale of two cities.
of-mi-an-m is the Title key for Of mice and men.
ham--- is the Title key for Hamlet.

Source fields: 245 \$a,b
Construction: Remove any non-filing chars (as specified by the second indicator) from the start of the title. Remove apostrophes and replace all non-alphanumeric chars by a single space (so hyphenated words become 2 words). Use the leading chars from the first 4 words in the pattern 3,2,2,1 (or as many chars as available from as many words as available).

6.2 Fields for Keyword Searching

6.2.1 Author (AUTH : 1003)

Source fields: 100 \$a,b,c,d,e,g,j,k,q,u,4
110 \$a,b,c,d,e,g,k,n,u,4
111 \$a,c,d,e,g,n,q,u,4
400 \$a,b,c,d,e,g,k,u,4 [LOCAL to U.S.]
410 \$a,b,c,d,e,k,n,u [LOCAL to U.S.]
411 \$a,c,d,e,g,n,q,u,4 [LOCAL to U.S.]
700 \$a,b,c,d,e,g,j,k,q,u,4
705 \$a [OBSOLETE]
710 \$a,b,c,d,e,g,k,n,u,4
711 \$a,c,d,e,g,n,q,u,4
715 \$a [OBSOLETE]
720 \$a,e
800 \$a,b,c,d,e,g,j,k,q,u,4
810 \$a,b,c,d,e,g,k,n,u,4
811 \$a,c,d,e,g,n,q,u,4

6.2.2 Uniform Title (UTI : 6)

Source fields: 130 X? \$a,d,f,g,h,k,l,m,n,o,p,r,s,t
240 ?X \$a,d,f,g,h,k,l,m,n,o,p,r,s
243 ?X \$a,d,f,g,h,k,l,m,n,o,p,r,s
700 \$t,m,n,o,p,r,s
710 \$t,f,g,h,k,l,m,n,o,p,r,s
711 \$t,m,n,o,p,q,r,s
730 X? \$a,d,f,g,h,k,l,m,n,o,p,r,s,t
800 \$t,f,g,h,k,l,m,n,o,p,r,s,v
810 \$t,f,g,h,k,l,m,n,o,p,r,s,v
811 \$t,f,g,h,k,l,m,n,p,r,s,v
830 ?X \$a,d,f,g,h,k,l,m,n,o,p,r,s,t,v

6.2.3 Key Title (KEYT : 33)

Source fields: 222 ?X \$a,b
534 \$k

6.2.4 Title (TITL : 4)

Source fields: 210 \$a,b
211 ?X \$a [OBSOLETE]
212 \$a [OBSOLETE]
214 ?X \$a [OBSOLETE]
241 ?X \$a,h [OBSOLETE]
242 ?X \$a,b,h,n,p,y
245 ?X \$a,b,f,g,h,k,n,p,s
246 \$a,b,f,g,h,n,p
247 \$a,b,f,g,h,n,p
740 X? \$a,h,n,p

6.2.5 Series Title (SERT : 5)

Source fields: 400 \$t,n,p,v [LOCAL to U.S.]
410 \$t,d,f,g,k,l,n,p,v [LOCAL to U.S.]
411 \$t,p,v [LOCAL to U.S.]
440 ?X \$a,n,p
490 \$a,v
800 \$t,f,g,h,k,l,m,n,o,p,r,s,v
810 \$t,f,g,h,k,l,m,n,o,p,r,s,v
811 \$t,f,g,h,k,l,m,n,p,r,s,v
830 ?X \$a,d,f,g,h,k,l,m,n,o,p,r,s,t,v

6.2.6 Library of Congress Subject Heading (LCSH : 27)

Source fields: 600 \$a,b,c,d,e,f,g,h,j,k,l,m,n,o,p,q,r,s,t,u,v,x,y,z,2,4
610 \$a,b,c,d,e,f,g,h,k,l,m,n,o,p,r,s,t,u,v,x,y,z,2,4
611 \$a,c,d,e,f,g,h,k,l,n,p,q,s,t,u,v,x,y,z,2,4
630 \$a,d,f,g,h,k,l,m,n,o,p,r,s,t,v,x,y,z,2
650 \$a,b,c,d,e,v,x,y,z,2
651 \$a,v,x,y,z,2
655 \$a,b,c,v,x,y,z,2 (if subfield 2 contains "lcsch" or if
2nd indicator = 0)
Construction: Take data only if 2nd indicator = 0 for source field.

6.2.7 Non-LC Subject Heading (SUBJ : 6---)

Source fields: 083 [All subfields]
 600 \$a,b,c,d,e,f,g,h,j,k,l,m,n,o,p,q,r,s,t,u,v,x,y,z,2,4
 610 \$a,b,c,d,e,f,g,h,k,l,m,n,o,p,r,s,t,u,v,x,y,z,2,4
 611 \$a,c,d,e,f,g,h,k,l,n,p,q,s,t,u,v,x,y,z,2,4
 630 \$a,d,f,g,h,k,l,m,n,o,p,r,s,t,v,x,y,z,2
 650 \$a,b,c,d,e,v,x,y,z,2
 651 \$a,v,x,y,z,2
 653 \$a
 654 \$a,b,c,v,y,z,2
 655 \$a,b,c,v,x,y,z,2
 656 \$a,k, v,x,y,z,2
 657 \$a,v,x,y,z,2
 658 \$a,b,c,d,2
 695 [All subfields] [Imp; Sth]
 752 \$a,b,c,d

Construction: For source fields 600-651, take data only if 2nd indicator
 ≠ 0.

6.3 *Fields for Browsing Indexes*

If we want the ability to browse indexes then we must create some special fields explicitly. This is because the database software is designed primarily for keyword searching.

6.3.1 Author Browse (ABRO : 1003)

Source fields: As in 6.2.1

6.3.2 Title Browse (TBRO : 4)

Source fields: As in 6.2.2 to 6.2.5 combined, but omit all \$h subfields.

6.3.3 Subject Browse (SBRO : 21)

Source fields: As in 6.2.6 and 6.2.7 combined.

6.3.4 Name Browse (NBRO : 1002)

Source fields: 100 \$a,b,c,d,e,g,j,k,q,u,4
110 \$a,b,c,d,e,g,k,n,u,4
111 \$a,c,d,e,g,n,q,u,4
~~400 \$a,b,c,d,e,g,k,u,4 [LOCAL to U.S.]~~
~~410 \$a,b,c,d,e,k,n,u [LOCAL to U.S.]~~
~~411 \$a,c,d,e,g,n,q,u,4 [LOCAL to U.S.]~~
600 \$a,b,c,d,e,g,j,k,q,u,4
610 \$a,b,c,d,e,g,k,n,u,4
611 \$a,c,d,e,g,n,q,u,4
700 \$a,b,c,d,e,g,j,k,q,u,4
710 \$a,b,c,d,e,g,k,n,u,4
711 \$a,c,d,e,g,n,q,u,4
~~800 \$a,b,c,d,e,g,j,k,q,u,4~~
~~810 \$a,b,c,d,e,g,k,n,u,4~~
~~811 \$a,c,d,e,g,n,q,u,4~~

****7. Fields in Alternate Scripts**

Required for storage (and eventual retrieval) of data in alternate scripts. Assumptions: (1) multiscript records follow Model A (i.e., some data is duplicated in 880 fields)[†]; (2) the language of cataloguing is English; (3) regular fields contain Latin script and 880 fields contain alternate/vernacular script.

If a library were to send us multiscript records which follow Model B they could still be loaded on the new database, but, for example, vernacular script data in 245 would be indexed in the regular Title field (6.2.4). This is not ideal but we believe that the bulk of titles in vernacular script will be indexed in the proper field (7.2) because, as far as we know, the vast majority of multiscript records follow Model A.

Each 880 field begins with subfield \$6, and cp0-2 of that subfield contains the linking tag of the regular field in Latin script. For example, in the 880 field below, the linking tag is 245:

```
245 10 $6880-01$aAka to kuro.
880 10 $6245-01/$1$a[Title in Japanese script].
```

N.B.: We intend to explore the possibility of creating some fields for *browsing* data in alternate scripts.

7.1 Author in Alternate Script (XAU)

```
Source field: 880
Subfields:    $6100 $a,b,c,d,e,g,j,k,q,u,4
              $6110 $a,b,c,d,e,g,k,n,u,4
              $6111 $a,c,d,e,g,n,q,u,4
              $6700 $a,b,c,d,e,g,j,k,q,u,4
              $6710 $a,b,c,d,e,g,k,n,u,4
              $6711 $a,c,d,e,g,n,q,u,4
Construction: Take data from source field only if linking tag in subfield
              $6 is 100, 110, 111, 700, 710 or 711. The source subfields
              depend on the linking tag.
```

7.2 Title in Alternate Script (XTI)

```
Source field: 880
Subfields:    $6130 X? $a,d,f,g,h,k,l,m,n,o,p,r,s,t
              $6245 ?X $a,b,f,g,h,k,n,p,s
              $6246 $a,b,f,g,h,n,p
              $6247 $a,b,f,g,h,n,p
              $6740 X? $a,h,n,p
Construction: Take data from source field only if linking tag is 130,
              245, 246, 247 or 740. The source subfields depend on the
              linking tag. Remove non-filing chars as specified.
```

[†] For a full discussion about the two models for multiscript records, *See* MARC 21 Bibliographic Format, Appendix D.
26/04/2010 @ 10:34

7.3 Place of publication in Alternate Script (XPOP)

Required to support limiting a search by place of publication, distribution, and so on.

N.B.: The regular MARC field containing Latin script is indexed in 8.1 below.

Source field: 880 \$a
 Construction: Take data from source field only if linking tag is 260.

7.4 Name of Publisher in Alternate Script (XPUB)

Required to support limiting a search by name of publisher, distributor, and so on.

N.B.: The regular MARC field containing Latin script is indexed in 8.2 below.

Source field: 880 \$b
 Construction: Take data from source field only if linking tag is 260.

7.5 Series Title in Alternate Script (XSER)

Source field: 880
 Subfields: \$6440 ?X \$a,n,p
 \$6490 \$a,v
 \$6830 ?X \$a,d,f,g,h,k,l,m,n,o,p,r,s,t,v
 Construction: Take data from source field only if linking tag is 440, 490
 or 830. The source subfields depend on the linking tag.
 Remove non-filing chars as specified.

****8. Other Fields**

8.1 Place of publication (PPUB : 59)

Required to support limiting a search by place of publication, distribution, and so on.

Source field: 260 \$a
 Construction: Omit field if source field contains "S.l." or "s.l."

8.2 Name of publisher (PUBR : 1018)

Required to support limiting a search by name of publisher, distributor, and so on.

Source field: 260 \$b
 Construction: Omit field if source field contains "s.n."

8.3 Pagination (PAGE : 6---)

Required to support limiting a search by pagination in book cataloguing records. The source field contains the pagination and normally includes several sequences of pages or leaves.

Source fields: 300 \$a, but only for Books (i.e., MTYP = BK)
 Construction: Normalise data as follows:
 1. Replace each mark of punctuation with a #
 2. Insert # between "p" and a "number"
 3. Insert # between "ca" and an integer

A "number" may consist of an integer or a Roman numeral.

E.g., "8, vii, ca. 300, 73 p." is indexed as
 "8 vii ca 300 73 p"

N.B.: Normalisation procedure will not always work.

8.4 Notes (NOTE : 63)

Required to support keyword searching of the 5XX fields. Like the Basic Bibliographic Fields in Section 6, these fields are basically textual in nature.

Source fields: All 5XX fields.
 Construction: Replace each subfield delimiter/code pair with a single #.
 Then take data as is.

8.5 Uniform Resource Identifier (URI : 1209)

Required to support searching for Uniform Resource Identifiers which define the location and access method of electronic resources.

Source fields: 856 \$u
505 \$u
506 \$u
514 \$u
520 \$u
530 \$u
540 \$u
545 \$u
552 \$u
555 \$u
563 \$u
583 \$u

Construction: Take data as is. Special chars (those which are not on the list of indexed chars) are treated as if they were #.

E.g.,

<http://susdl.fcla.edu/cgi-bin/cgiwrap/~fdl/fdlcgi?FA00000011%2Fjpg>

is indexed as several words as follows:

http susdl.fcla.edu cgi-bin cgiwrap fdl fdlcgi FA00000011 2Fjpg

****9. Display Fields**

9.1 MARC Record (MREC)

Used for display in all interfaces. Consists of a block of text representing a bibliographic record in MARC format.

Source fields: All MARC fields.

Construction: The pre-processing program converts records from exchange format to blocks of text in MARC format.

9.2 Short Display (SDIS)

Could theoretically be used for display in all interfaces. For end-users searching via Z39.50, this field could be sent in a 9XX field which some clients can be configured to display. For end-users searching via telnet/SSH or a web interface, this field could be used for brief record display.

Source fields: As specified below.

Construction: The field consists of a single line of text and could be constructed from data elements in the MARC record and from CURL control data, as follows:

Title proper and subtitle (245 \$a, b)
 "/" Statement of responsibility, if present (245 \$c)
 "(" ISBN ")", if present (020 \$a)
 Edition statement, if present (250 \$a)
 Country of publication (008 cp15-17)
 Date of publication (008 cp7-10)
 Holding library/Record source (Code from table in Appendix 1)
 Provenance code (000 \$f)
 ":" Encoding level, if present (Leader cp17)
 Subject Heading Count (0 if no 6XX fields present)
 CURL Minimum Standard Flag (000 \$e)

e.g.,

Romeo and Juliet/William Shakespeare(0140620931)

enk 1994 Cam c:f 02

Appendix I

CURL Library Codes

Library Code	Library No.
abn	26
bir	13
bnb	99
bri	29
cam	01
dur	18
edi	02
gla	03
imp	09
kcl	12
lco	98
lee	04
liv	19
lon	06
lse	11
man	05
ncl	17
nlw	71
nls	70
ntm	17
oxf	07
sas	21
she	14
sth	20
tcd	08
ucl	10
war	15
wel	22
cax	27
ram	23
rcm	24
ncm	25
tem	28

Appendix II

Mapping Between CURL Indexes and Z39.50 Use Attributes

The mapping is shown in the table of contents as well as each field definition. The number following the abbreviation (for the Database Field Name) specifies the Z39.50 use attribute to which the index is mapped. A specification of "6---" means that we think a private use attribute from, say, the 6000 block, will be necessary. For the three Date Fields, relation attributes are specified as well. With the Fields for Derived Searching, structure attributes are specified also. We have not yet specified anything with respect to Fields in Alternate Scripts.

Additional Z39.50 Use Attributes

We will also be able to support some additional Z39.50 use attributes which map to multiple CURL indexes:

6---	<i>Author Title Date Key</i> (auth, titl, date) maps to ATK & SDAT combined.
6---	<i>All Title Indexes</i> maps to UTI, KEYT, TITL & SERT combined.
21	<i>All Subject Indexes</i> maps to LCSH & SUBJ combined.
1016/1017	<i>All Indexes</i> .

Special Z39.50 Use Attribute

One attribute, 5000, will be configured to accept telnet style queries. It will map to individual or multiple CURL indexes depending on the query.

Appendix III

Response to this document from Cambridge, Liverpool and Oxford

Cambridge Response

Generally: where you talk about the need to "remove" a character or characters, can we take it that this also means closing up what remains so that there aren't any blanks? (See, for example, p. 18, sect. 5.1)

p. ii

Use attribute for 1.1 could also be 12 perhaps?

Use attribute for 2.8 could also be 1083 perhaps?

Use attribute for 2.9 could also be 1084 perhaps?

Use attribute for 2.10 could also be 1053 perhaps?

p. iii

Use attribute for 6.2.7 could also be either 29 or 1080 perhaps?

General note re Z39.50 mapping

I've noted 20 places where Local Use attributes will be required. Even allowing for the fact that 8 of these are in section 1, and that some of these are used only for internal purposes and wouldn't be searched directly by an outside user, should this number of local attributes be a cause for concern? This is a brand new database - would be expect so many of the indexes we think are important to be ones for which there's no international Use attribute assigned? The most obvious impact is on cross-database searching by Z39.50. It's unlikely that another database will support the same search using the same local use attribute as CURL. So one of the benefits of Z39.50 is lost along the way.

p. 21, sect. 6.2.4.

Agree with Warwick re subfield h in a browsable index.

There are various other fields/subfields containing title-related data which could be considered for a keyword index but which, again, must be excluded from a browse equivalent. 76X-78X for example.

p. 22, sect. 6.2.7.

The inclusion of local field 695 in the spec prompts me to wonder whether there are other local fields which people know they use and which they feel ought to be indexed. (Of course, one library's use may conflict with another's use of the same tag, but that also applies to 695...)

p. 22, sect. 6.3.1.

Severe doubts about the wisdom/desirability of including 490 in a browsable index - there is no non-filing indicator, yet the series statement may well begin with an article. Result... er, well, less that pretty. The whole point about 490 is that the series is not traced (or is traced differently) - including in a *keyword* index may be justifiable on grounds of "resource discovery", but there are significant problems when it comes to a browsable index.

p. 24, sect. 7.2

Add some of 730 to this list? (I suspect there's a reason why it's not there.)

p. 28, sect. 9.2.

Comment from our music specialist that now may be the time to press for inclusion of the Uniform Title in brief display. "It usually gives better information that the descriptive title" is how it was put to me. (See also below re maximum number of characters.)

More generally, this is clearly a hangover from the "telnet" era. Instead of being included in a Full Record as 9XX data is one option, for Z39.50 use, to return a Brief record (modified version of this section) if requested, and then to send the Full record only if subsequently asked?

What's the maximum number of characters permitted in this display? Is the restriction to a single line significant (yes, you can display more entries per screen, but only by truncating or completely omitting useful information)? The more pre-ISBN-era data we load, the less satisfactory this display becomes, perhaps?

Hope this helps.

Hugh Taylor
Head of Cataloguing, Cambridge University Library
West Road, Cambridge CB3 9DR, England

email: jrht3@cam.ac.uk fax: +44 (0)1223 339973
phone: +44 (0)1223 333069 (with voicemail) or
phone: +44 (0)1223 333000 (ask for pager 036)

Liverpool Response

1. UDC. If I'm right, CURL will no longer index UDC. Does this mean i) that UDC numbers will no longer appear on / download with the record, or ii) will still be on the record, but will no longer be searchable? We are not concerned whether UDC is indexed for searching or not, but would like them to remain as part of the record.

2. Combined title searches. Although the 'extended title search' previously available combined perhaps too many fields, I think we would prefer to keep a combined series/main title option search, since titles are often indexed differently by different libraries (same element indexed by series title (440 etc.) or main title (245 etc., depending on preference between multi-volume records and individual records for each volume). So we would like the option proposed in 6.3.2., but would possibly prefer an option combining just 6.2.4 and 6.2.5.

Martin Kelleher
Senior Information Assistant
Bibliographic Services
University of Liverpool

Oxford Response

2.4, Encoding Level (ENCL). Does your proposal to change the code for full level from "#" to "f" mean that libraries deriving records from CURL will then have to change the code back to "#" before saving it to their local systems in order to make the record MARC 21 compliant? This will make more work for cataloguers, assuming that they remember to do it, so is not very satisfactory. But presumably this is unavoidable?

2.5, Category of Material (CATM). We are pleased to see that this field is included, as we are making increasing use of field 007 in our records.

2.14, Microform Flag (MFRM). Should this take into account the fact that microform publications of, for example, early printed books, may be described in field 533 of a bibliographic record for the original?

3.1, Original Local Control Number (CTRL). If this index is built from only the original local control numbers of CURL members' record, without the library code, this may result in more than one hit for the same original local control number.

6.2.4, Title [keyword] (TITL). Include 700 \$t, 710 \$t, 711 \$t? Also, would it be worth including titles (\$t) from linking entry fields 76X-78X?

6.2.5, Series Title (SERT). Does the inclusion of the \$v (Volume/sequential designation) data from the series fields in this keyword index mean that searches could be narrowed to a particular number in a series? If so, would the data be taken as is or is some normalisation envisaged? If the former, some inconsistent search results may arise.

6.2.7, Non-LC Subject Heading (SUBJ). Field 752 is included as a source field, but in certain types of cataloguing (e.g. antiquarian books) this field is used for a place of publication or printing (in hierarchical form). Should the inclusion of this field in the SUBJ index be based on the material type of the record?

6.3.4, Name Browse (NBRO). This name index may be more useful as a keyword rather than a browse index, from the point of view of record retrieval. A name keyword could be combined with a title keyword to locate a record for an item when it is uncertain whether the name in question would have been coded by the cataloguer as an author or a subject (e.g. art exhibition catalogues).

7.2, Title in Alternate Script (XTI) and 7.5, Series Title in Alternate Script (XSER). We have had second thoughts about the need to include 130 and 830 respectively in these indexes. As these are uniform titles whose form is controlled by (usually) the LCNAF or BLNAL, we think it likely that these UTs would appear on the item in the original script.

8.1, Place of Publication (PPUB). Would you consider including 752 (if material type = BK) as a source field, to pick up place of publication/printing in a standard form from records for antiquarian material? (Cf. 6.2.7 above.)

8.5, Uniform Resource Identifier (URI). Does the indexing of this field take into account the possible substitution in the URI of "~" (spacing tilde) with "%7E", "_" (spacing underscore) with "%5F" and a space with "%20"? The example given is not very clear, as it seems to have "%2F" towards the end of the URI which looks like it might represent a period (dot) introducing the file extension "jpg", but it also contains a spacing tilde. This could result in the same URI being indexed differently if has been input in several source records in different ways.

Thank you for taking on board the comments we made back in December. We note with satisfaction the inclusion of a subject heading flag (1.8), publisher number index (4.4) and standard identifier (4.6).

Best regards,
Alison.

Alison Felstead, Catalogue Support Services, Bodleian Library,
Oxford University, Oxford OX1 3BG
Tel: +44 1865 277138, Fax: +44 1865 277182