

Guidelines for exchange of danMARC2 records in MarcXchange

Guidelines for exchange of danMARC2 records in MarcXchange

Series:

danZIG-specification 11

Authors:

Edited by Tommy Schomacker

Published by:

Danish Agency for Libraries and Media
H.C. Andersens Boulevard 2
DK 1553 Copenhagen V

Phone: +45 33 73 33 73

Mail: post@bibliotekogmedier.dk

Homepage: <http://www.bibliotekogmedier.dk/>

Language:

English

Digital ISBN:

978-87-92681-34-8

Version / Edition:

21-12-2011

Updates for this publication can be seen at:

<http://biblstandard.dk/kat/marcxchangeguidelinesupdate.htm>

Table of contents

Table of contents.....	3
1. Introduction.....	4
2. Background.....	4
3. Scope.....	4
4. Guidelines.....	4
4.1 danMARC2 records serialized according to MarcXchange.....	4
4.2 Transport protocols.....	5
4.2.1 FTP.....	5
4.2.2 Z39.50.....	5
4.2.3 SRU.....	5
4.3 Webservices.....	5
5. Conformance.....	5
6. Links.....	5
Annex A. Conversion between ISO 2709 and MarcXchange.....	6
General.....	6
Character set.....	6
Conversion of one danMARC2 record from MarcXchange to ISO 2709.....	6
Conversion of one danMARC2 record from ISO 2709 to MarcXchange.....	7

1. Introduction

These guidelines are part of a series of specifications for technical interoperability between libraries and between libraries and partners called danZIG specifications. danZIG is a committee representing library system vendors, library technical interoperability consultants and national library services supplemented by library representatives put together by Danish Agency for Libraries and Media and chaired by the agency. The purpose of danZIG is to advise the governmental agency Danish Agency for Libraries and Media, which have the responsibility for the published specifications.

2. Background

For nearly forty years ISO 2709 (Information and documentation – Format for information exchange) has been the preferred option for exchange of bibliographic records – MARC records.

ISO 25577 MarcXchange was published 1. December 2008.

The MarcXchange schema is useable for all MARC formatted records. It may be thought of as the xml version of ISO 2709.

3. Scope

These guidelines cover basic rules for exchange of danMARC2 records serialized according to MarcXchange. It may be supplemented by individual agreements between parties.

Other formats e.g. MARC 21 and presentation formats, are not addressed.

4. Guidelines

4.1 danMARC2 records serialized according to MarcXchange

MarcXchange can be used as an alternative to ISO 2709 for exchange of danMARC2 records. The records may contain all danMARC2 fields, including fields for local use.

The character set used in danMARC2 records serialized according to MarcXchange is Unicode (UTF-8 or UTF-16).

Use of “format” attribute is mandatory. Value is danMARC2.

The “type” attribute is not used.

No assumption is made on the content of the leader element.

See annex A for description of conversion of danMARC2 records between MarcXchange and ISO 2709.

4.2 Transport protocols

4.2.1 FTP

Files with danMARC2 records serialized according to MarcXchange can be exchanged as alternative to files with danMARC2 records serialized according to ISO 2709.

4.2.2 Z39.50

These guidelines supplements the existing Z39.50 specification, see section 5.3 Present service in the danZIG Profile Specification 2007: <http://www.bs.dk/publikationer/andre/danzig/01/>.

MarcXchange serialized danMARC2 records can be requested by specifying:

- OID = 1.2.840.10003.5.109.10
- *Element Set Name*: danmarcxchange

Ref: Z39.50 Implementor Agreement – Requesting XML Records

<http://www.loc.gov/z3950/agency/proposals/resolution/request-xml-revised-2009.html>

The result set is 0, 1 or many MarcXchange documents each with a single record. The collection element is not used.

4.2.3 SRU

MarcXchange formatted danMARC2 records can be requested by specifying:

- RecordSchema= danmarcxchange

The result set is 0, 1 or many MarcXchange documents each with a single record. The collection element is not used.

4.3 Webservices

It is recommended to use MarcXchange for exchange of danMARC2 records by the means of webservices.

5. Conformance

The parties are suppliers of integrated library systems , national systems and other library data bases.

The parties report their level of conformance to these guidelines to Danish Agency for Libraries and Media.

The information is published on the web: <http://biblstandard.dk/kat/>

6. Links

MarcXchange schema: <http://www.loc.gov/standards/iso25577/>

danMARC2: <http://www.kat-format.dk/danMARC2/default.html>

Annex A. Conversion between ISO 2709 and MarcXchange

General

Any danMARC2 record serialized according to ISO 2709 can be re-serialized according to MarcXchange.

Any size limited danMARC2 record serialized according to MarcXchange can be re-serialized according to ISO 2709. The attribute “size limited” indicates that the size of the danMARC2 record does not extend the size limits of an ISO 2709 record.

The leader element is used for the ISO 2709 Record label. Part of the ISO 2709 Record Label (position 0-4 Record length and position 12-16 Base address of data), contains information, which is only meaningful for the ISO 2709 record. Other parts of the Record label are not relevant for danMARC2, and the rest can be calculated from information in specific data fields.

It is recommended always to calculate the ISO 2709 Record Label when converting from MarcXchange to ISO 2709.

Character set

The character set used in danMARC2 records serialized according to ISO 2709 is the danMARC2 character repertoire.

The character set used in danMARC2 records serialized according to MarcXchange is Unicode (UTF-8 or UTF-16).

The documentation for the danMARC2 character repertoire gives a description of conversion between the danMARC2 character repertoire and Unicode.

Conversion of one danMARC2 record from MarcXchange to ISO 2709

The ISO 2709 record is created from the MarcXchange record element by generating a number of fields corresponding to the data field elements.

Each field has two indicators, and a number of subfields preceded by a subfield identifier equal to the code attribute for the corresponding subfield element.

The subfields of the field are constructed by converting the content of the subfield elements from Unicode (UTF-8 or UTF-16) to the danMARC2 character repertoire.

Then the directory is constructed by calculating the length and position of the fields.

Finally the record label is constructed according to this schema:

- Calculate length of record and insert the result in octets 0 to 4 (Record length)
- Insert the content of 004*r in octet 5 (record status)
- Insert the content of 009*a in octet 6 (implementation code: material type)
- Select value a, h, m or p from the content of 008*t and 004*a and insert the result in octet 7 (implementation code: bibliographical category)
- Insert the content of 004*a in octet 8 (implementation code: record type)
- Insert value hex(20) in octet 9 (implementation code: undefined)
- Insert value 2 in octet 10 (indicator length)
- Insert value 2 in octet 11 (identifier length)

- Calculate the address of the first field and insert the result in octets 12 to 16 (Base address of data)
- Select value 0, 1, 3, 4, 7, 8, 9 from the content of 008*v and insert the result in octet 17 (Positions defined by user systems: Cataloguing level)
- Insert value hex(20) in octets 18 and 19 (Positions defined by user systems: undefined)
- Insert value 4 in octet 20 (length in octets of the length of field part of each entry in the directory)
- Insert value 5 in octet 21 (length in octets of the starting character position part of each entry in the directory)
- Insert value 0 in octet 22 (length in octets of the implementation defined part of each entry in the directory)
- Insert value hex(20) in octet 9 (reserved for future use)

Conversion of one danMARC2 record from ISO 2709 to MarcXchange

A valid xml document with one record element is created according to the following pseudo code.

```
insert_element(record,
  attribute[format="danMARC2"],
  content
  [insert_element(leader, content[any valid string1]),
  for every field do
  insert_element(datafield,
    attribute[tag=copy(2709_field_tag)],
    attribute[ind1=copy(2709_ind1)],
    attribute[ind2= copy(2709_ind2)],
    content
    [for every subfield do
    insert_element(subfield,
      attribute[code= copy(2709_subfield_identifier)],
      content[convert_to_unicode(2709_subfield)]
    )
    enddo
  ]
  )
  enddo
  ]
)
```

If the xml document has a collection element as root element, it may contain more than one record element, each corresponding to an ISO 2709 record.

¹ The validity is specified by the Schema, i.e. according to the regular expression:

"\d{5}\p{IsBasicLatin}\p{IsBasicLatin}{4}\d\d\d{5}\p{IsBasicLatin}{3}\d\d\d\p{IsBasicLatin}"