

JAPAN/MARC21 (XML)の MODS への変換実験

Experiments on Transformation of JAPAN/MARC21 (XML) to MODS

浅倉 秀三 *

ASAKURA Shuzo

本論文は、JAPAN/MARC21 (XML)を MODS 形式へ変換することについて述べる。

LC が公開している MODS 形式への変換 XSL スタイルシートを利用して、JAPAN/MARC21 (XML)を MODS 形式へ変換した。その結果から、その変換に問題は見られないと考えられる。しかし、JAPAN/MARC21 ではよく使われ、MARC21 ではほとんど使われない要素がある。これらの要素に対し、その XSL スタイルシートはどのように対応しているかなどについては今後調査すべき重要な課題である。

[キーワード] MODS, MARC21, MARCXML, JAPAN/MARC21, JAPAN/MARC21 (XML)

1. はじめに

アメリカ合衆国議会図書館 (LC: The Library of Congress) [1] は、MARC21 のレコードを XML 形式へ変換した。これを MARCXML とよんでいる [2]。LC は、この MARCXML を "Dublin Core" (DC) や "Metadata Object Description Schema" (MODS) [3] など種々な形式への相互変換のバスとして位置づけている (図 1) [4]。そして、MARCXML を DC, MODS, HTML などの形式へ変換する XSL スタイルシートを公開している [2]。

国立国会図書館 (NDL: National Diet Library) [5] は、JAPAN/MARC [6] を作成している。この JAPAN/MARC のフォーマット (フィールド名、

サブフィールド名、コードなど) は、JAPAN/MARC 独自のものであった。筆者は、その JAPAN/MARC を MARCXML の XML 構造で XML 形式へ変換した (以後、これを JAPAN/MARC (XML) とよぶ) [7]。また、DC 形式への変換を行った。その後、筆者はこの JAPAN/MARC (XML) を MODS 形式へ変換することを研究した [8-10]。そこでは、筆者は、MARC21 の要素と JAPAN/MARC の要素との対応表を作ることをかなりの期間検討した。しかし、検討を重ねれば重ねるほど、MARC21 の全要素と JAPAN/MARC の全要素との対応表を作ることが不可能だと考えるようになった。そこで、文献 [10] のように MODS の要素と JAPAN/MARC の要素との対応表を作った。

* 中部大学 経営情報学部 経営情報学科

College of Business Administration and Information Science, Chubu University
E-mail: asakura@isc.chubu.ac.jp

2012年1月、JAPAN/MARCはMARC21フォーマットを採用した（以後、これをJAPAN/MARC21とよぶ）[11-13]。これは、変更というより新規格ともいえる。JAPAN/MARC21のフォーマットとMARC21のフォーマットとが同じになったことから、文献[8-10]における要素の対応表などは無用となった。

注 JAPAN/MARC21の要素は、MARC21の要素の部分集合である。

筆者は、国立国会図書館からダウンロードしたJAPAN/MARC21（2709）の書誌データをMARXMLへ変換するプログラムを開発した[14]。このMARXMLへ変換したJAPAN/MARC21をJAPAN/MARC21（XML）とよぶ。

また、JAPAN/MARC21フォーマットとLCMARC21フォーマットの同一性の検証（リーダの全項目とDCの処理対象要素）を行い、同一であることを確かめた[15]。

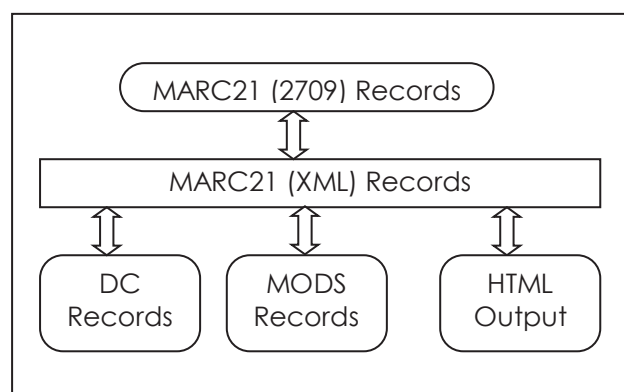


図1 MARXML Architecture [4]

本研究は、文献[8-10]に替えるものである。

本論文では、LCが公開しているMODS形式への変換XSLスタイルシート[2]を利用して、JAPAN/MARC21（XML）をMODS形式へ変換する。

2. MODSとMARC

MODSは、MARC21の書誌レコードから必要な要素を選択し、XML構造として、転記するために開発された。MODSは、MARC21の要素の部分集合からなり、要素名は数字ではなく、語句を用いている[3]。

1) MODS ファイルの構造

MODS ファイルの構造を次に述べる（図2）。

MODS ファイルは、<modsCollection>で始まり、</modsCollection>で終わる。

レコードは<mods>で始まり、</mods>で終わる。

レコードは、上位要素"Top Level Elements"からなる。上位要素には、"Subelements"（以後、これを中位または下位要素ともよぶ）を持つものもある。

```

<modsCollection>
  <mods>
    Top Level Elements
    Subelements
    ...
    Subelements
    Subelements
    Top Level Elements
  </mods>
  ...
  <mods>
    Top Level Elements
    Subelements
  </mods>
</modsCollection>
  
```

図2 MODSの構造

2) MODS レコード要素

MODSにおける上位要素は、MODS 3.5では、表1や付表1に示すように20である[16]。

表 1 MODS 3.5 上位要素の一覧

注 * 印は中位要素を持つ。

(01) titleInfo *	(11) note
(02) name *	(12) subject *
(03) typeOfResource	(13) classification
(04) genre	(14) relatedItem *
(05) originInfo *	(15) identifier
(06) language *	(16) location *
(07) physicalDescription *	(17) accessCondition
(08) abstract	(18) part *
(09) tableOfContents	(19) extension
(10) targetAudience	(20) recordInfo *

3. JAPAN/MARC21 (XML)から MODS へ

JAPAN/MARC21 (XML)に MARCXML 用スタイルシートを適用したとき、起こるエラーは次の二つに分類される。

- (1) 表示すべきでない要素を表示する。
- (2) 表示すべき要素を表示しない。

筆者は、MODS 3.5 で用いられるすべての要素 [17] が JAPAN/MARC21 に存在するかを調査した。その調査結果の一部を表 2 に示す。表中の◎印は JAPAN/MARC21 にある要素、*印はない要素を表す。これらの結果から、JAPAN/MARC21 の要素は、MODS 3.5 で用いられた要素の部分集合であることが明らかとなった。したがって、「(1) 表示すべきでない要素を表示する。」は起こらない。

表 2 MODS 3.5 要素と JAPAN/MARC21 要素との対応表の一部<titleInfo>

注 ◎印は JAPAN/MARC21 にある要素、* 印はない要素を示す。

MODS 3.5 要素	JAPAN/MARC21 要素
<titleInfo>	
<title>	◎245\$a*\$fgks
<subTitle>	◎245\$b
<partNumber>	◎245\$n*\$fgk
<partName>	◎245\$p*\$fgk
<nonSort>	◎245
<title>	◎210\$a
<subTitle>	◎210\$b
	*210\$2
<title>	*242
<subTitle>	*242
<partNumber>	*242
<partName>	*242
<title>	◎246\$a
<subTitle>	◎246\$b
	◎246\$i
<partNumber>	◎246\$n
<partName>	◎246\$p
<title>	◎246\$a*\$f
<subTitle>	◎246\$b
	◎246\$i
<partNumber>	◎246\$n
<partName>	◎246\$p
	*130, 240
<title>	*730
<partNumber>	*130, 240, 730
<partName>	*130, 240, 730
	*130, 240, 730
	◎710\$a*700, 710\$t, 711
<title>	◎740\$a*740\$h
<partNumber>	*740\$n
<partName>	*740\$p
	◎880\$6

4. 実験

実験の手順は次のとおりである。

- 1) 国立国会図書館から書誌データを ISO 2709 形式でダウンロードする。JAPAN/MARC21 (2709) である。
- 2) JAPAN/MARC21 (XML)形式に変換する。

3) LC が公開している MODS 形式への変換 XSL スタイルシート [2] を利用して MODS 形式へ変換する。

図 3 に MODS 3.4 形式への変換スタイルシートの一部を示す。

図 4 に JAPAN/MARC21 (XML) (JP98003970) を MODS 形式へ変換した結果の一部を示す。

図 4 の結果から、LC が公開している MODS 形式への変換 XSL スタイルシートを利用して、JAPAN/MARC21 (XML) を MODS 形式へ変換することに問題は見られないと考えられる。

しかし、JAPAN/MARC21 ではよく使われ、MARC21 ではほとんど使われない要素がある。これらの要素に対し、その XSL スタイルシートはどのように対応しているかなどについては今後調査すべき重要な課題である。

```
<!-- titleInfo -->
<xsl:for-each select="marc:datafield[@tag='245']">
  <xsl:call-template name="createTitleInfoFrom245"/>
</xsl:for-each>
<xsl:for-each select="marc:datafield[@tag='210']">
  <xsl:call-template name="createTitleInfoFrom210"/>
</xsl:for-each>
<xsl:for-each select="marc:datafield[@tag='246']">
  <xsl:call-template name="createTitleInfoFrom246"/>
</xsl:for-each>
<xsl:for-each select="marc:datafield[@tag='240']">
  <xsl:call-template name="createTitleInfoFrom240"/>
</xsl:for-each>
<xsl:for-each select="marc:datafield[@tag='740']">
  <xsl:call-template name="createTitleInfoFrom740"/>
</xsl:for-each>
<xsl:for-each select="marc:datafield[@tag='130']">
  <xsl:call-template name="createTitleInfoFrom130"/>
</xsl:for-each>
<xsl:for-each select="marc:datafield[@tag='730']">
  <xsl:call-template name="createTitleInfoFrom730"/>
</xsl:for-each>
```

図 3 の 1 MODS 3.4 形式への変換スタイルシートの一部 [2]

```
<xsl:for-each select="marc:datafield[@tag='242']">
  <titleInfo type="translated">
    <xsl:for-each select="marc:subfield[@code='y']">
      <xsl:attribute name="lang">
        <xsl:value-of select="text()"/>
      </xsl:attribute>
    </xsl:for-each>
    <xsl:variable name="title">
      <xsl:call-template name="chopPunctuation">
        <xsl:with-param name="chopString">
          <xsl:call-template name="subfieldSelect">
            <xsl:with-param name="codes">a</xsl:with-param>
          </xsl:call-template>
        </xsl:with-param>
      </xsl:call-template>
    </xsl:variable>
    <xsl:variable name="titleChop">
      <xsl:call-template name="chopPunctuation">
        <xsl:with-param name="chopString">
          <xsl:value-of select="$title"/>
        </xsl:with-param>
      </xsl:call-template>
    </xsl:variable>
    <xsl:choose>
      <xsl:when test="@ind2>0">
        <nonSort>
          <xsl:value-of select="substring($titleChop,1,@ind2)"/>
        </nonSort>
        <title>
          <xsl:value-of select="substring($titleChop,@ind2+1)"/>
        </title>
      </xsl:when>
      <xsl:otherwise>
        <title>
          <xsl:value-of select="$titleChop"/>
        </title>
      </xsl:otherwise>
    </xsl:choose>
    <title>
      <xsl:call-template name="chopPunctuation">
        <xsl:with-param name="chopString">
          <xsl:call-template name="subfieldSelect">
            <xsl:with-param name="codes">a</xsl:with-param>
          </xsl:call-template>
        </xsl:with-param>
      </xsl:call-template>
    </title>
    <xsl:call-template name="subtitle"/>
    <xsl:call-template name="part"/>
  </titleInfo>
</xsl:for-each>
```

図 3 の 2 MODS 3.4 形式への変換スタイルシートの一部 [2]

注 現在、MODS の版は仕様書などでは 3.5 [3] になっているが、XSL スタイルシートは 3.4 [2] である。

```

<mods version="3.4">
  <titleInfo altRepGroup="01" script="Latn">
    <title>夏目漱石作品集</title>
    <partNumber>第 1 巻</partNumber>
  </titleInfo>
  <name type="personal" altRepGroup="02" script="Latn">
    <namePart>夏目, 漱石</namePart>
    <namePart type="date">1867-1916</namePart>
  </name>
  <typeOfResource>text</typeOfResource>
  <originInfo script="Latn">
    <place>
      <placeTerm type="code" authority="marccountry">ja</pl
aceTerm>
    </place>
    <place>
      <placeTerm type="text">東京</placeTerm>
    </place>
    <publisher>昭和出版社</publisher>
    <dateIssued>1962.3</dateIssued>
    <dateIssued encoding="marc">1962</dateIssued>
    <issuance>monographic</issuance>
  </originInfo>
  <language>
    <languageTerm authority="iso639-2b" type="code">jpn</l
anguageTerm>
  </language>
  <physicalDescription>
    <form authority="marcform">print</form><form authorit
y="marccategory">text</form><form authority="marcsmd"
>regular print</form><extent>315p ; 19cm.</extent></ph
ysicalDescription>
    <tableOfContents>吾輩は猫である.</tableOfContents>
    <targetAudience authority="marctarget">general</targetAu
dience>
    <note type="statement of responsibility" altRepGroup="00
" script="Latn">夏目漱石 著.</note>
    <classification authority="kktb">KH426</classification>
    <classification authority="njb/09">918.68</classification>
    <titleInfo altRepGroup="01" script="CJK">
      <title>ナツメ ソウセキ サクヒンシュウ</title>
      <partNumber>1</partNumber>
    </titleInfo>
    <note type="statement of responsibility" altRepGroup="00
" script="CJK"></note>
    <titleInfo altRepGroup="01" script="Latn">
      <title>Natsume soseki sakuhinshu</title>
      <partNumber>1</partNumber>
    </titleInfo>
    ...

```

図 4 JAPAN/MARC21 (XML) (JP98003970)を MODS 形式へ変換した結果の一部

5. おわりに

筆者は、LC が公開している MODS 形式への変換 XSL スタイルシートを利用して、JAPAN/MARC21 (XML)を MODS 形式へ変換してみた。

その結果、その XSL スタイルシートを利用して MODS 形式へ変換することに問題は見られないと考えられる。

今後の課題は、種々な書誌レコードを用いて問題がないかを調査することである。

しかし、JAPAN/MARC21 ではよく使われ、MARC21 ではほとんど使われない要素がある。これらの要素に対し、その XSL スタイルシートはどのように対応しているかなどについては今後調査すべき重要な課題である。これは、「(2) 表示すべき要素を表示しない。」というエラーを起こさないためでもある。

参考文献

- [1] Library of Congress. "Library of Congress Home". <http://www.loc.gov/>, (accessed 2014-1-6).
- [2] Library of Congress. "MARC 21 XML Schema". <http://www.loc.gov/standards/marcxml/>, (accessed 2014-1-6).
- [3] Library of Congress. "Metadata Object Description Schema". <http://www.loc.gov/standards/mods/>, (accessed 2014-1-6).
- [4] Library of Congress. "MARC XML Architecture". <http://www.loc.gov/standards/marcxml/marcxml-architecture.html>, (accessed 2014-1-6).

[5] 国立国会図書館. “国立国会図書館—National Diet Library”. (オンライン), 入手先<<http://www.ndl.go.jp/>>, (参照 2014-1-6) .

[6] 国立国会図書館. “JAPAN/MARC マニュアル 単行・逐次刊行資料編 第3版(2009 フォーマット)” . (オンライン), 入手先<http://www.ndl.go.jp/jp/library/data/jmarc2009_ms_manual/manual.pdf>, (参照 2014-1-6) .

[7] 浅倉秀三. 日本語 MARC の XML データベースへの変換に関する一実験 — JAPAN/MARC と TRC MARC の変換 —. 日本図書館情報学会誌. Vol.50, No.4, 2004.12, p.159-170.

[8] 浅倉秀三. MODS 要素と JAPAN/MARC タグの対応に関する考察 — Top Level Elements の概要 —. 情報科学リサーチジャーナル (中部大学情報科学研究所) . Vol.15, 2008.3, p.93-100.

[9] 浅倉秀三. JAPAN/MARC と MODS の対応に関する研究 — MODS の上位要素の前半分 —. 情報科学リサーチジャーナル (中部大学情報科学研究所) . Vol.16, 2009.3, p.93-98.

[10] 浅倉秀三. JAPAN/MARC と MODS の対応に関する研究— MODS の全上位要素 —. 情報科学リサーチジャーナル (中部大学情報科学研究所) . Vol.17, 2010.3, p.3-11.

[11] 国立国会図書館. “JAPAN/MARC (M) (S) MARC21 フォーマット” . (オンライン), 入手先<http://www.ndl.go.jp/jp/library/data/pdf/JM_ms_MARC21revision.pdf>, (参照 2014-1-6) .

[12] 国立国会図書館. “JAPAN/MARC MARC21 フォーマット マニュアル 単行・逐次刊行資料編 2012” . (オンライン), 入手先<http://www.ndl.go.jp/jp/library/data/pdf/JAPANMARC_MARC21manual_MS.pdf>, (参照 2014-1-6) .

[13] 国立国会図書館. “書誌データのプロダクト

提供サービスが変わります: NDL 書誌情報ニュースレター2010年1号(通号12号)” . (オンライン), 入手先<http://www.ndl.go.jp/jp/library/data/bib_newsletter/2010_1/article_01.html>, (参照 2014-1-6) .

[14] 浅倉秀三. JAPAN/MARC MARC21 から MARCXML への変換プログラムの作成. 情報科学リサーチジャーナル (中部大学情報科学研究所) . Vol.19, 2012.03, p.87-92

[15] 浅倉秀三. JAPAN/MARC21 フォーマットと LC MARC21 フォーマットの同一性の検証 — リーダの全項目と Dublin Core の処理対象要素 —. 情報科学リサーチジャーナル (中部大学情報科学研究所) . Vol.20, 2013.03, p.87-96

[16] Library of Congress. "Outline of Elements and Attributes in MODS Version 3.5". <http://www.loc.gov/standards/mods/mods-outline-3-5.html>, (accessed 2014-1-6).

[17] Library of Congress. "MARC 21 to MODS 3.5 Mapping ". <http://www.loc.gov/standards/mods/mods-mapping.html>, (accessed 2014-1-6).



浅倉秀三
経営情報学部
経営情報学科
教授 工学博士
専門分野:
自然言語文字列処理,
電子図書館

付表 1 MODS 3.5 要素の一覧

注 星印は、MODS 3.5 で追加された要素を示す。↓ ↑ は、MODS 3.5 で順位が変更になった要素を示す。

(1) titleInfo	(12-1) topic	(14-19) extension
(1-1) title	(12-2) geographic	(14-20) recordInfo
(1-2) subTitle	(12-3) temporal	(15) classification
(1-3) partNumber	(12-4) titleInfo	(16) location
(1-4) partName	(12-5) name	(16-1) physicalLocation
(1-5) nonSort	(12-6) geographicCode	(16-2) shelfLocator ↓
(2) name	(12-7) genre	(16-3) url ↑
(2-1) namePart	(12-8) hierarchicalGeographic	(16-4) holdingSimple
(2-2) displayForm	(12-8-1) continent	(16-4-1) copyInformation
(2-3) affiliation	(12-8-2) country	(16-4-1-1) form
(2-4) role	(12-8-3) province	(16-4-1-2) sublocation
(2-4-1) roleTerm	(12-8-4) region	(16-4-1-3) shelfLocator
(2-5) description	(12-8-5) state	(16-4-1-4) electronicLocator
(2-6) etal*	(12-8-6) territory	(16-4-1-5) note
(3) typeOfResource	(12-8-7) county	(16-4-1-6)
(4) genre	(12-8-8) city	enumerationAndChronology
(5) originInfo	(12-8-9) island	(16-5) holdingExternal
(5-1) place	(12-8-10) area	(17) accessCondition
(5-1-1) placeTerm	(12-8-11) extraterrestrialArea	(18) part
(5-2) publisher	(12-8-12) citySection	(18-1) detail
(5-3) dateIssued	(12-9) cartographics	(18-1-1) number
(5-4) dateCreated	(12-9-1) scale	(18-1-2) caption
(5-5) dateCaptured	(12-9-2) projection	(18-1-3) title
(5-6) dateValid	(12-9-3) coordinates	(18-2) extent
(5-7) dateModified	(12-10) occupation	(18-2-1) start
(5-8) copyrightDate	(13) classification	(18-2-2) end
(5-9) dateOther	(14) relatedItem	(18-2-3) total
(5-10) edition	(14-1) titleInfo	(18-2-4) list
(5-11) issuance	(14-2) name	(18-3) date
(5-12) frequency	(14-3) typeOfResource	(18-4) text
(6) language	(14-4) genre	(19) extension
(6-1) languageTerm	(14-5) originInfo	(20) recordInfo
(6-2) scriptTerm*	(14-6) language	(20-1) recordContentSource
(7) physicalDescription	(14-7) physicalDescription	(20-2) recordCreationDate
(7-1) form	(14-8) abstract	(20-3) recordChangeDate
(7-2) reformattingQuality	(14-9) tableOfContents	(20-4) recordIdentifier
(7-3) internetMediaType	(14-10) targetAudience	(20-5) recordOrigin
(7-4) extent	(14-11) note	(20-6) languageOfCataloging
(7-5) digitalOrigin	(14-12) subject	(20-6-1) languageTerm
(7-6) note	(14-13) classification	(20-7) descriptionStandard
(8) abstract	(14-14) relatedItem	
(9) tableOfContents	(14-15) identifier	
(10) targetAudience	(14-16) location	
(11) note	(14-17) accessCondition	
(12) subject	(14-18) part	