Metadata formats and the new markup language from library field

Marianna Salgáné Medveczki

Abstract. Using new communication technologies libraries must face continuously new questions, possibilities and expectations. This study discusses some library-related aspects of our electronic era and how computer-related data formats affect bibliographic dataprocessing to give a survey of the most important results.

Altough library systems have been supporting the creation and maintenance of MARC records for decades, the need to create other forms of metadata is relatively new. In higher education it is important that students get acquainted with these new aspects of bibliographic description and bibliographic dataprocessing.

In our department I launched the BDML development project in order to standardize bibliographic description (based on Hungarian standard 3424/1) with the help of XML. The development aims at the creation of a set of standards that provides information about the structure of bibliographic description in XML form. In the second part of the paper I give a brief presentation and summary of BDML.

Key words and phrases: computer-related bibliographic data formats, metadata formats, XML, BDML.

ZDM Subject Classification: B20.

1. Introduction

Librarians have been engaged in resource description for as long as there have been libraries. Using new communication technologies libraries must face continuously new questions, possibilities and expectations in this field.

Today libraries work with this new infrastructure induces many challenges. One of the most significant challenges is moving from a relatively homogenous bibliographic environment to a diverse one. Despite these challenges such changes are achievable and necessary to exploit possibilities of new metadata and technologies like the internet and XML.

In our department I launched the BDML development project in order to standardize bibliographic description with the help of XML. The development aims at the creation of a set of standards that provides information about the structure of bibliographic description in XML form. I hope that our DTD for a bibliographic element set can be used for library applications and tasks in other fields too.

2. The results of electronization to bibliographic dataprocessing

The first bibliographic formats for the exchange of bibliographic and related information in the machine-readable form between different types of computer systems were created more than 30 years ago. The Library of Congress created MARC (Machine Readable Cataloguing) bibliographic format, on which other versions of MARC format were developed to meet the needs of a variety of countries and bibliographic systems, such as: the USMARC in the USA, the CANMARC in Canada, from which MARC21 format was developed in 1999, the UKMARC in the United Kingdom, the UNIMARC (Universal MARC), slightly adapted for the use in the European countries, etc. Based on the UNIMARC format, some national bibliographic formats have been created [3].

The cataloguing rules may differ for different MARC formats: ISBD, AACR, LCSH.

The evolution of information technologies leads to the improvement of computer systems. In addition to the development of computers and media types internet has a great influence on data structure as well. Since the introduction of MARC bibliographic format, technology of data exchange between computers and between different computer systems has reached a very sophisticated stage and has contributed to the creation of new standards in this field.

The ISO 2709 standard specifies the method of transport of MARC bibliographic record fields on the magnetic tape. ISO 2709 only specifies the function of bibliographic record transport on to the magnetic tape that is used for the implementation of MARC bibliographic format in certain computer systems. The

content, length and structure of individual fields, subfields and indicators are determined by the MARC bibliographic format itself. The ISO 2709 standard describes a generalized structure, a framework designated especially for communication between data processing systems and not for the use of a processing format within systems. This is emphasized on the grounds of the structure of ISO records being complex and inadequate for data processing. It is not only unreadable for human, but requires development of very complex programs for record content processing.

The evolution of information technologies leads to the improvement of computer systems. In addition to improved computers, the abilities of data transfer among computers and computer systems have also become better. New media types have been developed. The development of computers, media types and internet has a great influence on data structure, as well. Nowadays,, data exchange is carried out over the internet to a large extent and not so much by physical media.

In the mid 1990s, the Network Development and MARC Standards Office developed two SGML DTDs that supported the conversion of data from MARC data structure to SGML (and back) without loss of data. The project of creating DTD specifications and conversion utilities was opened for input from any interested MARC and/or SGML users. The alpha version of the MARC DTDs was made available in May 1996. In 1997, the work on the MARC-to-SGML conversion utilities began.

In 2001, with XML specification development, the SGML DTD was fully converted to XML DTD, which determines particularly the class of XML documents representing MARC records.

In 2002, in the Library of Congress was developed an XML schema for a bibliographic element set that could be used for a variety of purposes and defined the new XML format of bibliographic records, referred as MODS. It includes a subset of MARC fields and is more user-oriented than MARC XML format [20].

3. Metadata formats

Altough library systems have been supporting the creation and maintenance of MARC records for decades, the need to create other forms of metadata is relatively new.

Altough metadata is a topic of major interest in library and information science, the term itself comes from the field of computer science.

Metadata is data used to describe other data. Specifically, the term refers to data used to identify, describe, or locate information resources, whether these resources are physical or electronic. While structured metadata processed by computers is relatively new, the basic concept of metadata has been used for many years in helping manage and use large collections of information. Library card catalogs are a familiar example of such metadata.

By the early 1990s, the term metadata was being used in the sense of the information necessary to make computer files useful to humans, particularly in relation to scientific, social science, and geospatial datasets. One of the first specifications to call itself metadata was the FGDC's Content Standard for Digital Geospatial Metadata, version 1 of which was issued in 1994.

Library historians are likely to see the 1990s as a decade of particular excitement, creativity and change. It will certainly be known for the rise of the web, and as the decade that the digital library was invented. It may also be known for an almost explosive proliferation of metadata schemes. The first draft version of the TEI Guidelines, including the definition of the TEI header, was distributed in 1990. The first version of the FGDC Content Standard for Digital Geospatial Metadata was released in 1994. The workshop that drafted the original Dublin Core Metadata Element Set was held in 1995. The alpha version of the EAD was released in 1996. The Core Categories for Visual Resources version 2.0 was released by the VRA in 1997. The Data Documentation Initiative was established in 1995 and released an XML version of the DDI metadata standard for social science data resources in 1997. The learning resources community produced both the Dublin Core-based GEM element set in 1998 and the IMS Meta-data Specification in 1999. And so on; this list is only a sampling. "In the metadata garden, truly a thousand flowers are blooming." (P. Caplan, [5])

With the rise of internet computing and the web, the term metadata began to be used in the context of describing information objects on the network. The term entered the working vocabulary of mainstream librarianship around 1995 with the creation and promotion of the Dublin Core Metadata Element Set. Librarians were quick to realize that they had been creating data about data, in the form of cataloging, for centuries. However, there is inconsistent use of the term metadata even within the library community, some using it to refer to the description of both digital and nondigital resources, and others restricting it to electronic resources only. Altough the more restrictive interpretation is probably closer to the original computer science concept, it is certainly more useful to think of metadata as descriptive of all types of information resources, including print publications.

Many collections consist of both digital and nondigital objects, and even digital objects often trace their ancestry to an artifactual original.

Another often assumed constraint is that metadata itself must be electronic, regardless of the nature of the object described. In practical terms this is not terribly restrictive, as most resource description today is created and stored in digital form.

As a final variation on the term, the W3C appears to have adopted the most restrictive definition of all: "Metadata is machine understandable information for the web" [18]. The requirement that metadata be machine-understandable is almost the exact inverse of the original need expressed by scientists and social scientists dealing with numeric datasets, who needed metadata to make their computer data human-understandable. This definition also expressed a more subtle requirement that metadata is "for the web", disqualifying not only the card and book catalogs allowed by publishers, but also disqualifying resource description accessible via all other internet protocols and all non-web-based computer systems.

It should be abundantly clear by now that there is no right or wrong interpretation of metadata, but that anyone using the term should be aware that it may be understood differently depending on the community and context within which it is used.

In library and information science a fairly liberal definition is employed. Metadata is here used to mean structured information about an information resource of any media type or format. It does place two constraints on what qualifies as metadata. First, the information must be structured, which is to say it cannot be a randomly accumulated or represented set of data elements, but must be recorded in accordance with some documented metadata scheme. Second, the metadata must describe an information resource.

Similarly, the UKOLN says that metadata "is normally understood to mean structured data about data digital (and non digital) resources that can be used to help support a wide range of operations. These might include, for example, resource description and discovery, the management of information resources (including rights management) and their long-term preservation" [17].

Overall, the most useful discussions of metadata are not concerned with what it applies to but rather with what it is intended to accomplish.

This has been a mixed blessing for libraries, presenting (as most innovations do) both opportunities and challenges. On the positive side, it has given us new options for describing materials that are poorly served by the MARC suite of

standards, and it has created a renewed aspect in resource description. At the same time, these new formats have placed new burdens on the library profession. There are mature, well-developed tools for creating and managing traditional cataloging.

It could be argued that beyond being intended for electronic description of information resources, the metadata schemes have little in common. They have different intended users and different intended uses. They are to varying degrees "bibliographic", in terms of being designed to support the FRBR (a report published by IFLA in 1998). Some are defined in terms of DTDs, while others are semantic categories independent of syntax.

In fact, metadata can be used to describe many different types or levels of entity, from abstract concepts to physical objects. One of the fundamental aspects of definition for a metadata scheme or element is specifying which type(s) of entities the scheme or element can apply to.

4. About the XML in few words

With the development of the internet the XML has also been developed and is one of the means for stopping the internet chaos.

XML is an open standard, a universal language for data on the web. XML is nearly six-years-old standard designed for the description and computer-based management of (semi)-structured data and structured texts. XML gives developers the power to deliver structured data from a wide variety of applications and it is also an ideal format for server to server transfer of structured data.

XML also isn't limited for internet use and is an especially valuable tool in the field of library. In fact, XML's main strength – organizing information – makes it perfect for exchanging data between different systems. Tools that work with the XML can be used to process XML records without incurring additional costs associated with one's own software development. In addition, XML is also a suitable format for library web services.

XML is a metalanguage based on SGML standard for specifying new formal languages and XML permits a high level of customization of tags (using the XML namespace mechanism). XML can be tought of as a subset of SGML designed with tighter rules, fewer features, and fewer options, in order to make it easier to process. XML is more suitable for use on the web. For example, XML tag names are case sensitive and can contain non-ASCII UNICODE characters. This

makes XML adapted for an international environment in which usage of the roman alphabet cannot be assumed.

W3C created XML to cover a wide area of electronic publishing. But, the right use of XML is attained in the exchange of many different types of data over the internet.

The development of XML and related specifications is an active initiative of the W3C. Related specifications approved or under development include an XML linking language (XLink, XPointer), an XML query language (XQuery), a stylesheet specification (XSL/XSLT, XPath). An important related activity is the development of XML Schema, an alternative way of defining document types.

DTDs, which are also used in SGML, define the structure of XML documents with a list of legal elements. It is easiest to think of a DTD is a context-free grammar. Grammar of XML defines XML elements and attributes and their combinations. DTD specification defines grammar with paticular syntax rules for descriptions of elements and attributes of languages, their values and possible combinations, repeatability, obligation, etc. Every DTD specification defines one class of XML documents. DTD not enables defining of complex data types, this is the disadvantage of DTD concerning XML schema. A well-formed XML document conforms to the XML syntax, and a well-formed XML document that conforms to its DTD is called valid.

The DOM is an application programming interface (API) for well-formed XML documents. It defines the logical structure of documents and the way a document is accessed and manipulated. In the DOM specification, the term "document" is used in the broad sense - increasingly, XML is being used as a way of representing many different kinds of information that may be stored in diverse systems, and much of this would traditionally be seen as data rather than as documents. Nevertheless, XML presents this data as documents, and the DOM may be used to manage this data.

DOM Core specification defines a set of objects and interfaces for accessing and manipulating document objects. The functionality specified (the Core functionality) is sufficient to allow software developers and web script authors to access and manipulate parsed XML content inside conforming products. Using the DOM, programmers can build documents, navigate their structure, and add, modify, or delete elements and content.

As a W3C specification, one important objective for the DOM is to provide a standard programming interface that can be used in a wide variety of environments and applications. The DOM is designed to be used with any programming language.

Since the introduction of XML specification, the definition of the bibliographic record format has experienced a new dimension. XML is not only a new bibliographic record format: the appropriate specification of XML bibliographic records in bibliographic information systems. For implementation of functions, such as validation of records or their transformation to different structures, many ready tools have been developed in Java and most of them declared as open source.

5. About the BDML

BDML is markup language in library enviroment creating bibliographic description for different library materials and for representing information in the web. BDML is based on Hungarian cataloguing rules. BDML is expressed using the XML DTD of the W3C. It serves well as syntax for metadata.

The BDML DTD specifies an order and grouping of elements. These are the internal structures of the DTD.

On the other hand BDML is an experiment to show how we can create a markup language for bibliographic data which can perform the following expectations and demands for traditional and electronic documents in library environment:

- better utilization of possibilities of bibliographic description in electronic environment preserving the identification function,

- flexible, enlargable, platform-independent metalanguage or dataformat (content-based – and not form-based – description of datastructure which is stored electronically in understandable from for both computer programs and human beings),

- separate metadata and output format,

- self-descriptive data-elements,

- flexible join to document (XML-based languages are enlargable because one can declare at will its own elements),

- fit to bibliographic data in traditional and electronic documents and describe theirs relations, correlations,

- in the bibliographic dataformats critical demand is the ordering about sequence of elements,

- give sequence-rules depending on context and occurence-rules relation to the obligatory and repeatable elements,

- fully separate of content of elements and punctuation symbols of bibliographic description,

- string bibliographic records in hypertext system.

BDML is a descriptive metadata system. Feature of descriptive metadata is understood to serve the purposes of discovery (how one finds a resource), identification (how a resource can be distinguished from other, similar resources), and selection (how to determine that a resource fills a particular need). Descriptive metadata may also be used for collocation (bringing together all versions of a work) and aquisition (obtaining a copy of the resource, or access to one). (Traditional library cataloging viewed as metadata is primarily descriptive.)

BDML gives a data structure and not a data content standard. It does not prescribe how one formulates the data that appears in any given data element – that is role of external national (as Hungarian standards) or international (as ISBDs) data content standards.

There is a possibility to interlock the elements in arbitrary depth in XML and in this way high flexibility is available.

BDML uses language-based tags. Using the XML DTD, BDML defines main elements, child elements (i.e. subelements), and attributes of elements and subelements. Content of elements are included in the lowest level elements to avoid "mixed content", occurs when some elements contain character data interspersed with child elements. For instance, if `<titles>` contains subelements for `<titleproper>`, `<subtitle>`, `<othertitle>`, then `<titles>` is only a wrapper tag to include the more specific elements `<titleproper>`, `<subtitle>`, `<othertitle>` and does not contain any character data. (A "wrapper" tag is one that it used only as an element that binds together child elements, but contains no other data than tags.)

One of design principles is that the order of elements within a BDML instance is relevant to the output of that data. This structure facilitates preparation of different ouput sequence or product.

The transformation of these records to several meta formats is simple (for instance with XSLT), records in this format are human readable and end user oriented, processing of records in this format produces analytical output and

validation in a simple way, and programs for record processing in this format are not so complicated as programs for working with records in ISO 2079 format.

An "XSLT stylesheet" may be written to transform the BDML data in some way for output. Examples include using a stylesheet to place the record into a template with easy-to-understand element names in XML; using a stylesheet to formulate a display that looks like a catalog card; using a stylesheet to transform coded data into textual form.

As a conclusion, short foretaste of BDML DTD was extracted; this transplants the monographic description of books into XML-language.

```
<?xml version = "1.0" encoding = "UTF-16LE"?>
<!-- XML BIBLIOGRAPHIC DESCRIPTION DTD
     Bibliographic Description Markup Language(BDML)
     AUTHOR: Medveczki M. (marianna@inf.unideb.hu)
     DOCUMENT TYPE: MONOGRAPHIC PUBLICATIONS (based on Hungarian standard
     3424/1)
     VERSION: 2.0 (2005.03.08.) -->

<!-- MAIN STRUCTURE -->

<!-- Bibliographic description is document element -->
<!ELEMENT mp_bibl_desc (single_level | multi_level | component_parts)>
<!ATTLIST mp_bibl_desc id ID #REQUIRED autopsy (y | n) "y">

<!ELEMENT single_level (single | single_volume)>

<!ELEMENT single (area1, area2?, area3, area4, area5?, area6, area7)>
```

The `<single>` element contains the monographic description of books (single-part publications). It contains seven subsections:

- `<area1>` information about the title and those responsible for the intellectual content;
  ```
  <!ENTITY % repeated_first_area "(point, first_area)*,
  (colon, subtitle)*, (colon, othertitle)*">
  <!-- Publication can contain individual works by different authors,
  unless a linking word or phrase is given in the publication -->
  <!ELEMENT area1 (newindent, char_position3., first_area,
  %repeated_first_area;)>
  ```
- `<area2>` information relating to the edition of the book;
  ```
  <!ELEMENT area2 (pointspacedashspace, edition_statement,
  %opt_statements_of_responsibility;)>
  ```
- `<area3>` information on the publication details of the book;

```
<!ELEMENT area3 (pointspacedashspace, %publisher;, %dates;,
%printer;)>
```

- `<area4>` information about the physical extent of the book;

```
<!ELEMENT area4 (pointspacedashspace, extent, (colon, illustration)?,
semicolon, dimensions, (plus, accompanying_material)?)>
```

- `<area5>` information about the series to which the book belongs;

```
<!ELEMENT area5 (pointspacedashspace, ocurvedbracket, series,
ccurvedbracket, (point, ocurvedbracket, series, ccurvedbracket)*)>
```

- `<area6>` notes providing additional information about the text;

```
<!ELEMENT area6 ((pointspacedashspace, note, (pointspacedashspace,
note)*) | ((newindent, char_position3.), note, ((newindent,
char_position3.), note)*))>
```

- `<area7>` information relating to the standard number (ISBN) and terms of availability of the book.

```
<!ELEMENT area7 ((pointspacedashspace, elements, (point,
elements)*) | ((newindent, char_position3.), elements,
((newindent, char_position3.), elements)*))>
```

Marking elements of convential signs (prescribed punctuation) of bibliographic description were defined as empty elements, i.e. these elements haven't subject matter.

```
<!-- Definition of Punctuation Symbols -->

<!ELEMENT ccurvedbracket EMPTY>
<!ELEMENT char_position3. EMPTY>
<!ELEMENT colon EMPTY>
<!ELEMENT comma EMPTY>
<!ELEMENT csquarebracket EMPTY>
<!ELEMENT diagonal_slash EMPTY>
<!ELEMENT equals_sign EMPTY>
<!ELEMENT exlam_mark EMPTY>
<!ELEMENT newindent EMPTY>
<!ELEMENT ocurvedbracket EMPTY>
<!ELEMENT osquarebracket EMPTY>
<!ELEMENT plus EMPTY>
<!ELEMENT point EMPTY>
<!ELEMENT pointspacedashspace EMPTY>
<!ELEMENT question_mark EMPTY>
<!ELEMENT semicolon EMPTY>
```

Some of the most important parts of DTD was represented also in treeview form at the end of the paper (see you Figure 1 – Figure 4 in the Appendix).

6. Role of different metadata (e.g. BDML) in higher education of librarians

Libraries face challenges in integrating descriptive metadata for different documents and resources with traditional cataloguing data. In our days librarians work mostly in a MARC-based environment, all the same it must be thinking about how MARC data can and will be used within a broader metadata environment.

Higher education in field of library and information science is in the process of being transformed to reflect new trends, such as the increasing importance of developing a library metadata management design.

Our students have a working knowledge of ISBDs, Hungarian standards, online cataloguing (in integrated library systems) and MARC on courses of Bibliographic Dataprocessing. We discuss among others catalog management defined as the continuous upgrading and updating of a catalog, regardless of its format, be it card, book or online. At the same time it is important that students will be able to understand machine-assisted generation of cataloguing and have instruction in bibliographic control practices that will help them continue to play a significant role in shaping library services in the emerging digital information environment.

The focus, however, has generally been across-the-board education of librarians, with the hope of preparing professionals who have the background and conceptual understanding to deal with information in all its forms and the tools and technologies necessary to manage its acquisition, maintenance, and use. These are laudable goals and accomplishments, even if they are not particularly new.

Hence special additional courses are necessary to cover overview of basic concepts of 21st century bibliographic control, including specific metadata standards and applications, to cover overview of the scope and nature of the planned changes in cataloguing rules and potential impacts on cataloguing practices, policies, organisations, and user services. These cources will give an introduction to different metadata too (what is it, why it was developed, its potential uses, advantages features) and how it is being used in various library projects.

In addition students can latch on to different projects of defining the vocabulary of metadata and describing the important components of the metadata environment in a clear and articulate manner.

It is important to understand that this approach does not add requirements for all students; it simply provides an opportunity for further advanced study.

7. Summary

While cataloguing principles have shown remarkable persistence in the web environment, cataloguing practices are becoming increasingly isolated due to their dependence on MARC coding. It is necessary to explore balancing valuable traditions with new technologies to create a potential foundation for future access to information in a distributed digital environment.

An alternative can be an experimental DTD, which enables bibliographic records to be created, encoded and styled using XML, which may predominant web environment of the future. The BDML DTD is based not on MARC but on the structure of ISBDs and Hungarian cataloguing rules. These experiments can provide a useful indication of ways in which XML capabilities – particularly the distinction between visual format and semantic structure – can be used enhance the intellectual activity of cataloguing in an environment that permits far more interoperability with other web resources and tools.

Explanation of words

AACR	Anglo-American Cataloguing Rules
BDML	Bibliographic Description Markup Language
DOM	Document Object Model
DTD	Document Type Description
EAD	Encoded Archival Description
FGDC	Federal Geographic Data Committee
FRBR	Functional Requirements for Bibliographic Records
GEM	Gateway to Educational Materials
IFLA	International Federation of Library Associations and Institutions
ISBD	International Standard Bibliographic Description
ISBN	International Standard Book Number
LCSH	Library of Congress Subject Headings
MARC	Machine Readable Cataloguing
MODS	Metadata Object Description Schema
SGML	Standard Generalized Markup Language
TEI	Text Encoding Initiative
UKOLN	U.K. Office for Library and Information Networking
XML	Extensible Markup Language

XSLT	Extensible Stylesheet Language Tranformation
VRA	Visual Resources Association
W3C	World Wide Web Consortium

Appendix
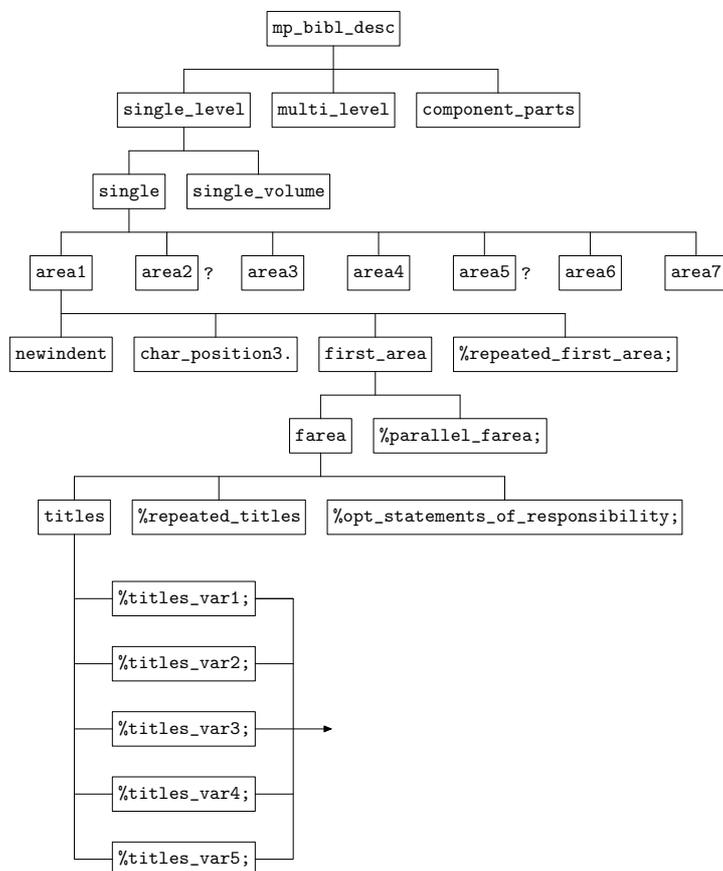
Figure 1. Graphical representation of Title and Statements of Responsibility Area based on BDML DTD

Metadata formats and the new markup language from library field   99

%titles_var1

titleproper   *   equal_sign   ptitleproper   *   colon   subtitle   *   colon   othertitle

%titles_var2

titleproper   *   equal_sign   ptitleproper   +   colon   subtitle   %pothertitles;

%titles_var3

titleproper   %psubtitles;   *   colon   othertitle

%titles_var4

titleproper   +   %ptitlesubtitle;   *   colon   subtitle   colon   othertitle

%titles_var5

%titlesubtitleothertitle;   %ptitlesubtitleothertitle;

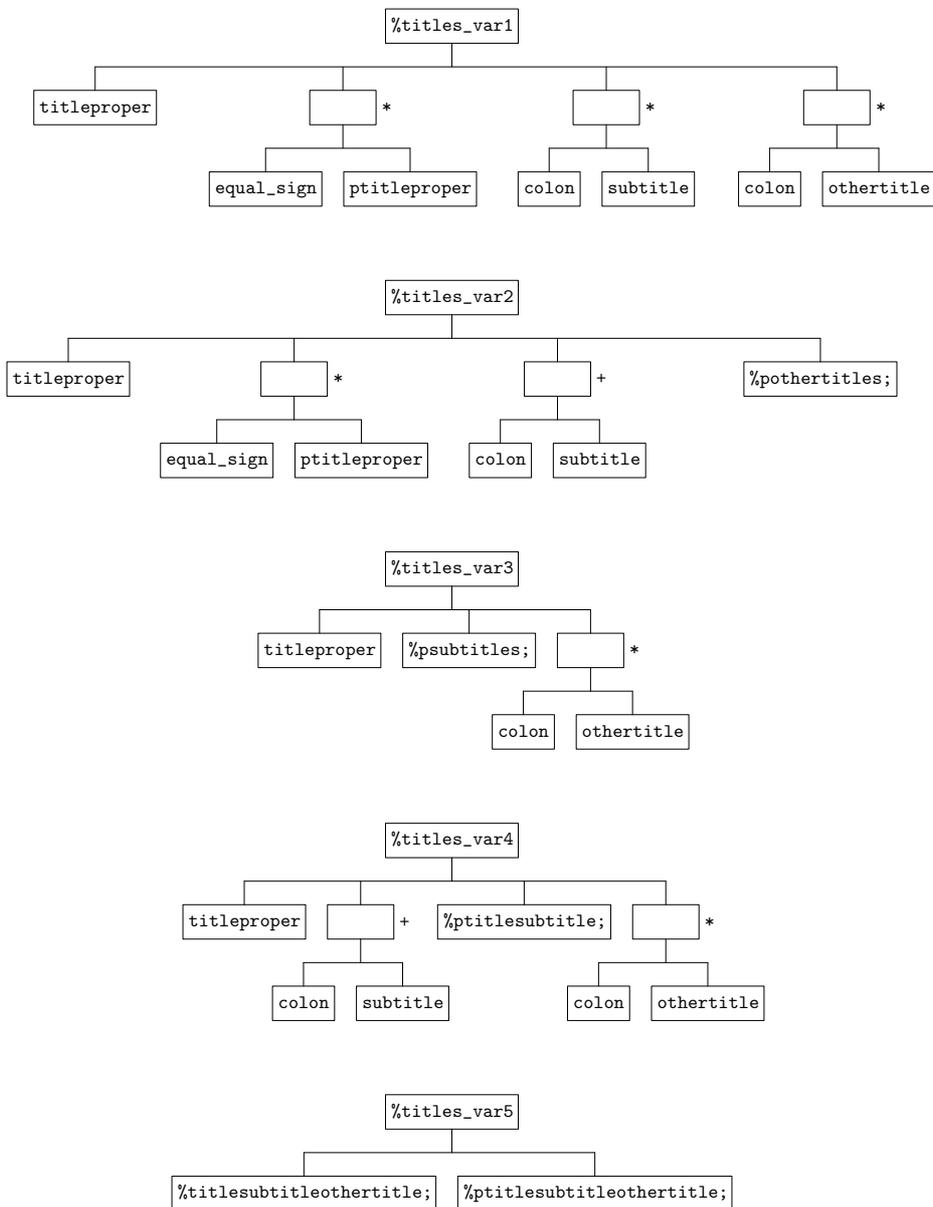*Figure 2.* Graphical representation of *titles* element

```
<area1>
<newindent/>
<car_position3./>
<first_area>
    <farea>
        <titles>
            <titleproper>
                <title>Dühöngő ifjúság</title>
            </titleproper>
            <colon/>
            <subtitle>színmű három felvonásban</subtitle>
            <equal_sign/>
            <ptitleproper>
                <ptitle>Look back in anger</ptitle>
            </ptitleproper>
            <colon/>
            <psubtitle>a play in three acts</psubtitle>
        </titles>
        <diagonal_slash/>
        <statements_of_responsibility>
            <statement_of_responsibility>
                <statresp>John Osborne</statresp>
            </statement_of_responsibility>
            <semicolon/>
            <statement_of_responsibility>
                <osquarebracket/>
                <statresp.os>ford. Ottlik Géza</statresp.os>
                <csquarebracket/>
            </statement_of_responsibility>
        </statements_of_responsibility>
    </farea>
</first_area>
</area1>
```

Figure 3. Simple example of an occurence of the first area

```
245 10 $aDühöngő ifjúság : $bSzínmű három felvonásban = Look back in
       anger : A play in three acts /$cJohn Osborne ; [ford. Ottlik
       Géza]
```

Figure 4. Previous example is in USMARC format (the subtitle, the
parallel title and the parallel subtitle are mixed up in subfield b; sim-
ilarly, the statements of resposibility combine in subfield c; and all
subfields contain not only contents but punctuation symbols too)

References

[1] Berke Barnabásné, A nemzeti bibliográfiai ajánlások megújítása, *Könyv, könyvtár, könyvtáros, 9. évf.* **8** (2000), 11–20, Párizs 1977 – Koppenhága 1998.

[2] M. Buckland, *A manifesto*, Redesigning library services, Chicago, ALA, 1992.X.

[3] G. Budimir, MARC records and XML, *Infoteka* **5**, no. 1–2 (2004), 121–136.

[4] M. Burkart, *Dokumentationssprachen*, In: Buder, M. – Rehfeld, W. – Seeger, T. (hrsg.), Grundlagen der praktischen Information und Dokumentation, München, 1990, K. G. Saur..

[5] P. Caplan, International Metadata Initiative,
`http://www.loc.gov/catdir/bibcontrol/caplan_paper.html`.

[6] Document Object Model Core.,
`http://www.w3.org/TR/2004/REC-DOM-Level-3-Core-20040407/`.

[7] Dublin Core Metadata Initiative (DC). DCMI, 1995–2003, `http://dublincore.org/`.

[8] Extensible Markup Language (XML) 1.0. (2. ed.). W3C Recommendation, 2000,
`http://www.w3.org/TR/REC-xml`.

[9] Extensible Stylesheet Language (XSL), Version 1.0. W3C Recommendation, 2001,
`http://www.w3c.org/TR/xsl`.

[10] Family of ISBDs: Publication list, International Federation of Library Associations and Institusions, 2003, `http://www.ifla.org/VI/3/nd1/isbdlist.htm`.

[11] R. Hagler, *The bibliographic record and information technology*, 3. ed., Chicago – London, ALA-Ottawa, CLA, 1997.

[12] T. Horváth and I. Papp (szerk.), *Könyvtárosok kézikönyve*, 2.köt., Osiris, Budapest, 2001.

[13] IFLA study on functional requirements for bibliographic records, *International Cataloguing and Bibliographic Control* **24**, no. 3 (1995), 43–50.

[14] Information Retrieval (Z39.50): Application Service Definition and Protocol Specification, National Information Standards Organization, 2002,
`http://www.niso.org/standards/resources/Z39-50-200x.pdf`.

[15] Mandatory data elements for internationally shared resource authority records, Report of the IFLA UBCIM Working Group on Minimal Level Authority Records and the ISADN, 1998.

[16] MARC Standards. Network Developement and MARC STANDARD Office, Library of Congress, 09/16/2003, `http://lcweb.loc.gov/marc/`.

[17] Metadata (UKOLN website), `http://www.ukoln.ac.uk/metadata/`.

[18] Metadata and Resource Description, `http://www.w3.org/Metadata`.

[19] Metadata Encoding & Transmission Standard,
`http://www.loc.gov/standards/mets/`.

[20] Metadata Object Description Schema (MODS), Library of Congress, 11/25/2003,
`http://www.loc.gov/standards/mods/v3/mods-3-0.xsd`.

[21] Dick R. Miller, *XML and MARC: a choice or replacement?*, ALA., Chicago, 2000, 12, `http://elane.stanford.edu/laneauth/ALAChicago2000.html`.

[22] Network Development and MARC Standards Office,
`http://www.loc.gov/marc/ndmso.html`.

[23] ONIX for Books. EDItEUR, `http://www.editeur.org/onix.html`.

[24] International Standard ISO 8879 Information Processing – Text and Office Systems – Standardized Generalized Markup Language (SGML), First Edition, International Organization for Standardization, 1996.

[25] Resource Description Framework (RDF), `http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/`.

[26] Roy Tennant, Building a new bibliographic infrastructure, Digital Libraries, *Library Journal* (1/15/2004), 38.

[27] Roy Tennant, MARC must die, *Library Journal* **10** (2002), `http://libraryjournal.reviewsnews.com`.

[28] Roy Tennant, Metadata leadership, Digital Libraries, *Library Journal*, (8/15/2004).

[29] UNIMARC Manual: Bibliographic Format, International Federation of Library Associations and Institusions, 1994, `http://www.ifla.org/VI/3/p1996-1/sec-uni.htm`.

[30] UNIMARC Manual: Bibliographic Format, APPENDIX H: cataloguing Rules and Formats Codes, International Federation of Library Associations and Institusions, 1994, `http://www.ifla.org/VI/3/p1996-1/appx-h.htm`.

[31] XML Path Language, W3C Recommendation, 1999, `http://www.w3.org/TR/xpath`.

[32] XSL Transformations (XSLT), Version 1.0. W3C Recommendation, 1999, `http://www.w3c.org/TR/xslt`.

[33] Erik Vajda, A besorolási adatok egységesítése, Fogalmak tisztázása és terminológiai javaslat, *Tudományos és Műszaki Tájékoztatás* 37 évf. **5** (1990), 186–190.

[34] Web Services Activity, W3C Recommendation, 2002–2003, `http://www.w3.org/2002/ws/`.

MARIANNA SALGÁNÉ MEDVECZKI
FACULTY OF INFORMATICS
UNIVERSITY OF DEBRECEN
H–4010 DEBRECEN P.O. BOX 12
HUNGARY

E-mail: `marianna@inf.unideb.hu`