XML BASED FORMAT FOR THE EXCHANGE AND THE STORAGE OF BIBLIOGRAPHIC RECORDS

Dimitri Busch (1)
(1) Fraunhofer-Informationszentrum,Raum und Bau IRB, Nobelstrasse 12, 70569 Stuttgart
(Germany), dbu@irb.fraunhofer.de

1. INTRODUCTION

Bibliographic references play an important role in the knowledge management, especially in research institutions [9]. The purpose of the article is to describe a XML based approach to the exchange and the storage of bibliographic records by the example of the International Construction Database (ICONDA). The ICONDA is a bibliographical database covering the worldwide literature on all fields of building construction, civil and construction engineering, architecture and town planning [6].

2. APPROACHES TO THE BIBLIOGRAPHIC METADATA EXCHANGE

The ICONDA Communication Format (ICF) is currently used for the information exchange between the ICONDA database, information providers and information users [2]. The ICF supports the ISO 2709 [5] standard that is also supported in other common exchange formats, e.g. in CCF [8] and MARC [1]. Although the ICF provides a fine bibliographic description of various literature types, it is difficult to create and to process the data because of the complicated record structure. The format also has some restrictions for multilingual literature titles. These problems can be solved by means of the XML format because the format is easy to read, can be processed by standard parsers and supports various character encoding systems. The existing approaches to the bibliographic metadata exchange with the XML can be devided into formats that emulate data structures of the ISO 2709 standard, e.g. MARCXML [7] and MarcXchange [4], and formats that are independent from the ISO 2709, e.g. MODS [1] and EXodic [3]. After the studies of these approches, we have developed the Bibliographical Metadata Format (BMF) that is based on the XML, and supports existing data elements from the ICF and ISO 2709. The BMF is covered below.

3. BIBLIOGRAPHICAL METADATA FORMAT (BMF)

The BMF uses the XML element *<record>* to describe a bibliographical record. The *<record>* element consists of a set of child *<datafield>* elements that decribe data fields. A data field is identified by the *tag* attribute. For example, the tag 200 identifies a data field that consists of a title, and the tag 300 identifies a field that consists of a person name. Each *<datafield>* element consists of one or more child *<subfield>* elements. A *<subfield>* element has a *code* attribute that identifies the subfield. For example, the subfield "A" of the data field 300 consists of a family name of a person. The field tags and subfield codes are compatible with the ICF. A BMF record can be logically devided into segments that are important to decribe publications that are parts of other publications. The segments are identified by the *segment* attribute of the *<datafield>* element.

A sample record that describes a book article is shown in Fig. 1. The record consists of two segments (0 and 1). The first segment (0) describes an article. The article has the title "Modern Architecture". The second segment (1) describes a book that consists of the article. The book has the title "Architectural Styles" and the author "Max Mustermann".

```
<record identifier="id001" status="a" level="a">
…..
        <datafield tag="200" segment="0" ind1="0" ind2="1">
                <subfield code="A">Modern Architecture</subfield>
        </datafield>
…..
        <datafield tag="200" segment="1" ind1="0" ind2="1">
                <subfield code="A">Architectural Styles</subfield>
        </datafield>
        <datafield tag="300" segment="1" ind1="0" ind2="0">
                <subfield code="A">Mustermann</subfield>
                <subfield code="B">Max</subfield>
                <subfield code="E">020</subfield>
        </datafield>
…..

</record>
```
Fig. 1: Elements of a sample record in the BMF

4. CONCLUSION

The BMF records can be easily created and processed because they are based on the XML. It is also not compicated to convert bibliographical records from the ICONDA Communication Format (ICF) into the BMF because the BMF supports existing data elements from the ICF.

The new format can be used not only to exchange bibliographical records, but also to store the records in the database. The newly reimplemented ICONDA database is currently using the SQL Server 2005 to store bibliographical records in the BMF. It is also possible to execute the full text indexing and retrieval of the records.

REFERENCES

[1] Guenther, Rebecca, "New and traditional descriptive formats in the library environment". DC2004 IFLA session, 2004. 9 Mar. 2009 <dc2004.library.sh.cn/english/prog/ppt/ifla.ppt>

[2] ICONDA Agency, "ICONDA communication format: format for the exchange of records in the frame of the international construction database". Fraunhofer- Informationszentrum Raum und Bau (IRB). 5.ed. Stuttgart: IRB Verlag, 1990

[3] INIST, „EXodic - The INIST exchange format". INIST-Instutute for scientific and technical information. Nancy (France). 9 Mar. 2009 <http://international.inist.fr/article159.html>

[4] ISO 25577 Standard „Information and documentation. MarcXchange". International Organization for Standardization. Geneva: ISO, 2008

[5] ISO 2709 Standard "Information and documentation. Format for information interchange". International Organization for Standardization. 3.ed. Geneva: ISO, 1996

[6] International Construction Database (ICONDA). Fraunhofer- Informationszentrum Raum und Bau (IRB). Stuttgart. 9 Mar. 2009 <http://www.irbdirekt.de/iconda/

[7] The Library of Congress, "MARCXML. MARC 21 XML Schema", 2008. 9 Mar. 2009 <http://www.loc.gov/standards/marcxml/>

[8] UNESCO, "CCF: The Common Communication Format". Paris: UNESCO, 1988

[9] Wilde, Erik, "References as Knowledge Management". Swiss Federal Institute of Technology, 2004. 16 Mar. 2009, <http://www.istl.org/04-fall/article4.html>