

MARC Futures

**International Workshop: MARC 21 – Experiences,
Challenges, and Visions
May 2007**

**Sally H. McCallum
Library of Congress
smcc@loc.gov**

Introduction

❖ What is MARC?

- Collection of data elements?
- Technical encoding of data elements?
 - Structure (syntax)
 - Tagging (semantic)
- Rules for data content?

Outline

Nine format attributes of the “future” environment:

- ❖ XML
- ❖ Granularity
- ❖ Versatility
- ❖ Extensibility
- ❖ Hierarchy support
- ❖ Crosswalks
- ❖ Tools
- ❖ Cooperative management
- ❖ Pervasive

Citation

- ❖ **Beth Goldsmith and Frances Knudson**
 - ***The Search for a Common Standard for Digital Repository Metadata***
 - **D-Lib Magazine, September 2006, Vol.12, No.9**
- ❖ LANL needed a “vendor-neutral” format for 87,000,000 metadata records from a variety of sources.
- ❖ Article describes their choice of MARC21 in the MARCXML container and illustrates a forward-looking use of the format.
- ❖ www.dlib.org/dlib/september06/

XML syntax

- ❖ **MARC data element family** (examples)
 - MARCXML for MARC in XML
 - MODS

- ❖ **Future?**
 - MARCXML streamlined?
 - What is after XML?

XML examples

MARCXML Example:

```
<datafield tag="100" ind1="1" ind2=" " >  
  <subfield code="a">Tan, Chung Lee</subfield>  
  <subfield code="d">1949-</subfield>  
</datafield>
```

MODS Example:

```
<name type="personal">  
  <namePart>Tan, Chung Lee</namePart>  
  <namePart type="date">1949-</namePart>  
</name>
```

Granularity

- ❖ **Lots of detail in data elements**
 - Approximately 200 fields; 1800 subfields
- ❖ **Lots of parsing of data**
- ❖ **Lots of coded data**
- ❖ **Administrators do not like**

- ❖ **Take some out?** (statistics)
 - Difficulty obtaining agreement
 - Variety of material and user types
 - International value of coded data
 - Can data be removed, really?

Some statistics

- ❖ ***MARC Content Designation and Utilization Analysis***
(William Moen and Shawne Miksa)(www.mcd�.unt.edu)
 - **Looking at book and serial records**
 - **34,546,200 records from OCLC WorldCat**
 - **90% of occurrences: 28 fields**
 - **7,595,887 records from LC**
 - **90% of occurrences: 21 fields**
 - **Also studying non-book records and subfield usage**
 - **Caveats: analysis not complete, non-book fields need to be sorted out, etc.**

- ❖ **Results similar to a 1997 German study of field usage for LC records**

Versatility

- ❖ **Simple to complex, general to specialized description** (example)
- ❖ **MARC targets rule independence**
- ❖ **More community profiling in the future?**
 - **MARC lite and MODS are types of profiles**
 - **RDA profile; CCO profile?**
 - **FRBR “work” profile?**
 - **Profiles based on statistical studies?**

Simple or full examples

Brief record:

001 97052102

008 971216

100 1 \$aTan, Chung Lee\$d1949-

245 \$aFinland

Full record:

001 97052102

005 19971629094819.1

008 971216s1998 wiu 001 0 eng

020 \$a0836820134 (library binding)

040 \$aDLC\$cDLC\$dDLC

082 \$a394.2694897\$221

100 1 \$aTan, Chung Lee\$d1949-

245 10 \$aFinland\$b2by Tan Chung Lee

260 \$aMilwaukee, Wis.\$bGareth Press Publ.\$c1998

300 \$a178 p. \$bill.\$c21 cm

650 0 \$aFestivals\$zFinland

Extensibility

❖ Grows and adapts

- UNICODE adaptation
- URI links – digital resource, cover art, etc.
- Geographic coordinates for authority records

❖ Additions in the future?

- Rights information? – link to external rights record?
- Preservation information? – link to PREMIS data?
- Media specific technical data? – combine through METS?
- New user communities – cultural, museum, etc.

Hierarchy support

❖ MARC

- Fields and subfields
- Linking entry fields to other resources
- URLs to other records

❖ Other hierarchy solutions in the future?

- In MARCXML, marc:collections model
- In MODS, related item model
 - CD – cut
 - Journal issue – article – image in article
- RDF “flat” preference vs. XMLs love of hierarchy

Crosswalks

❖ MARC with other formats

- MARC ↔ MODS
- MARC ↔ Dublin Core (simple)
- MODS ↔ Dublin Core
- ONIX → MARC
- Unimarc → MARC
- Digital geospatial ↔ MARC

❖ Element sets and rules

- FRBR → MARC
- RDA → MARC
- CCO? Digital resource metadata?

❖ Where does this stop?

Tools

❖ Library of Congress

- MARC 21 \leftrightarrow MARCXML (example)
- XSLT transformations for many crosswalks

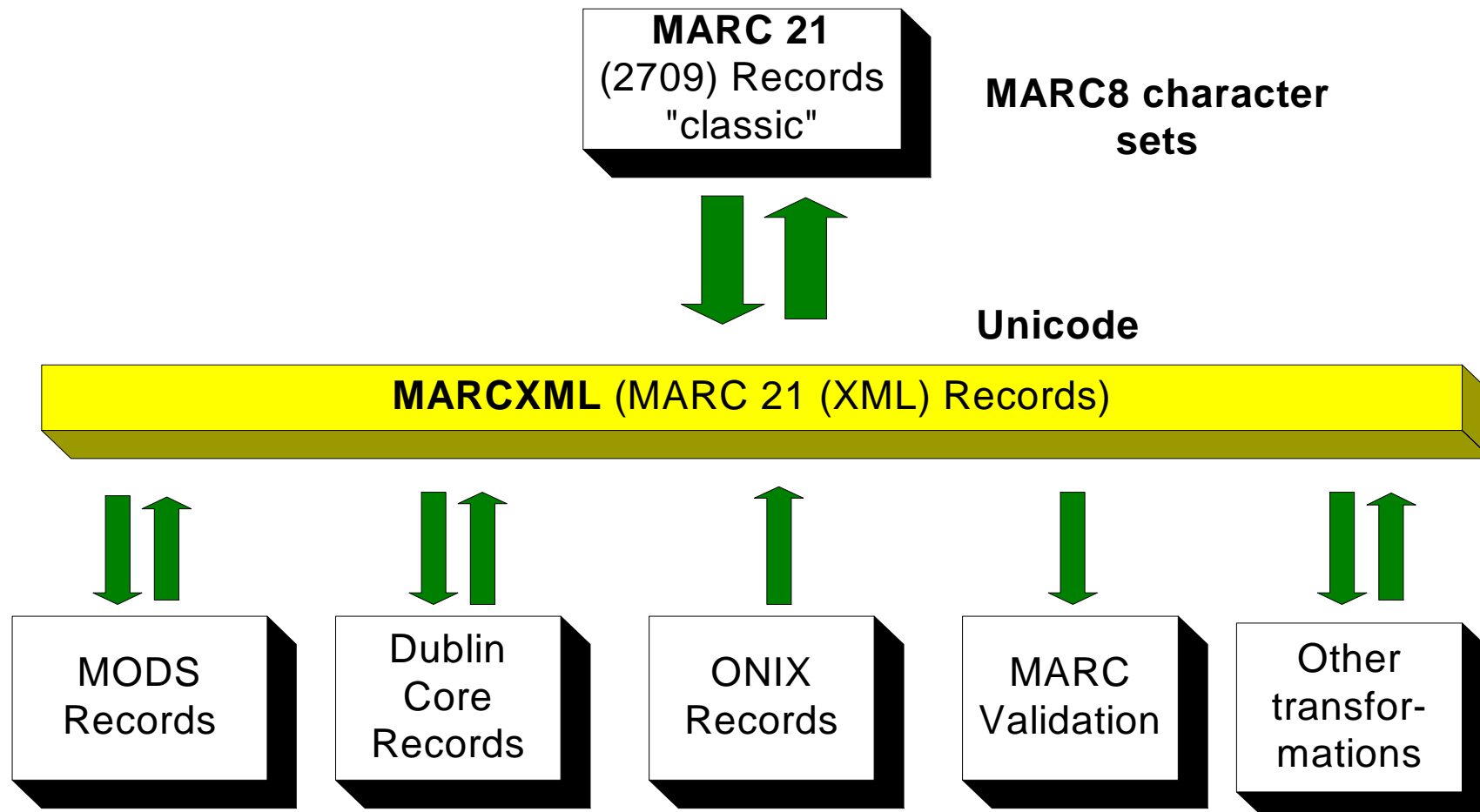
❖ Community developed

- MarcEdit, MARC Report, MARC Magician, XMARC, ZMARCO, USEMARCON from Europe, MARC4J from Netherlands, Aurora ZMarc Collector from Australia, etc.

❖ In the future more sharable *open source* software utilizing MARC in XML?

MARC tool kit

transformations (green arrows)



Cooperative management

- ❖ **National Library stakeholders**
- ❖ **National Committees like MARBI, CCM, BICBSG**
- ❖ **MARC Advisory Committee – includes national libraries, specialized material specialists (e.g., visual material, maps, etc.), system vendors, networks**
- ❖ **Listservs**
 - **MARC**
 - **Unicode-MARC**
- ❖ **Future?**

Pervasive

❖ Longevity

❖ Lingua franca

❖ Global use

- USA, Canada, Latin America, South Africa, etc.
- UK, Spain, Norway, Poland, Switzerland, Sweden, Czech Republic, Croatia, and others – Germany and Austria
- Australia, New Zealand, Hong Kong, China, Philippines, Thailand, Viet Nam, and others
- Over a billion MARC records in 1000s of systems

❖ Future?

- Experimentation by users with MARCXML
- MODS and XML open new use arenas?

In Summary

- ❖ **Pushing further into XML**
- ❖ **Continue broad coverage but with more profiling?**
- ❖ **More coordination with other format, element set, and rule developments**
- ❖ **Evolutionary**