The Crosswalk Web Service

Jean Godby
Research Scientist
OCLC Research
The Crosswalk Web Service at OCLC

- Enables OCLC to translate from one *metadata format* to another.
  - A “metadata format” is a triple that consists of a metadata schema, a structural encoding, and a character encoding.
  - Supported standards are bibliographic, but the software can handle other types of data.
- Can be called from any product or service that processes metadata.
- A version with a slightly different interface resides on the OCLC Enterprise Bus.
The translations

Inputs
- ONIX Books
- MARC XML
- MODS
- DC XML
- OAI-DC XML
- OCLC CDF

MARC 21-2709

OCLC’s Common Data Format

Outputs
- OCLC MARC
- OCLC CDF
- ONIX Books
- MARC XML
- DC XML
- DC-Qualified
- MODS
- ONIX Serials
Data flow for a single translation

MARC input

Convert to input structure

Translate to DC Terms

Convert to output structure

DC Terms output

<?xml version="1.0" encoding="UTF-8"?>
<qualifieddc xmlns:dcterms='purl.org;dc/terms'>
  <dctermsset>
    <dcterms:spatial>
      northwest
    </dcterms:spatial>
  </dctermsset>
</qualifieddc>
The MARC-Dublin Core relationship

Marc
- DC Simple
- OCLC Local
  - DC Simple

DC Terms
- DC Terms
  - DC Simple
  - OCLC Local
    - DC Terms
      - DC Simple
    - Other
      - local variants
<xml version="1.0" encoding="UTF-8" ?>
  <dctermsset>
      <dc:creator>International Conference on Dublin Core and Metadata Applications (2001)</dc:creator>
      <dc:contributor>Oyama, Keizo</dc:contributor>
      <dc:contributor>Gotoda, Hironobu</dc:contributor>
      <dc:contributor>Kokuritsu Jōhōgaku Kenkyūjo</dc:contributor>
      <dc:contributor>Dublin Core Metadata Initiative</dc:contributor>
      <dc:abstract>Includes bibliographical references.</dc:abstract>
      <dc:subject>Material support for the DC-2001 conference provided by: National Institute of Informatics ... [et al.; edited by ...</dc:subject>
      <dc:extent>297 p. : ill. ; 30 cm.</dc:extent>
      <dc:identifier>4924600989</dc:identifier>
      <dc:identifier>9784924600980</dc:identifier>
      <dc:language>English and Japanese</dc:language>
    </qualifieddc>
  </dctermsset>
</xml>
From ContentDM to WorldCat

- OAI Harvester
- DC-MARC
- CONTENTdm
- Custom maps
- Master file and derivatives linked in metadata
- Store digital master files in Digital Archive
- CONTENTdm copies digital master files to archival volumes
- Customized CONTENTdm web sites
- Build collections with CONTENTdm Acquisition station(s)
- CONTENTdm Collection Administrator
- CONTENTdm WorldCat Metadata Harvesting program
- OCLC Staff, data analysis
- Store digital master files in Digital Archive
Next Generation Cataloging

Vendor and publisher records

ONIX  ONIX  ONIX  ONIX

ESweep

Translate enriched records

WorldCat

To CDF

WorldCat

To ONIX

Vendors & publishers

NextGen process flow
A graphical user interface

Inputs

<map>
Source: MARC 245 $a
Target: ONIX Title
</map>

<map>
Source: MARC 650 $a
Target: ONIX Subject
</map>

Outputs

Editing interface

Search interface

Map database

Standard translation

Implied translation

Application profile

Version upgrade
For more information

- Research reports
  - Encoding Application Profiles in a Computational Model of the Crosswalk
  - Toward element-level interoperability in bibliographic metadata
  - A Repository of Metadata Crosswalks
  - Two Paths to Interoperable Metadata

- The public demo
  - OCLC Crosswalk Web Service Demo

- OCLC Information and Services for Publishers