New Orleans, LA • June 24, 2018

OCLC Research BIBFRAME 2.0 Converter Analysis

LC BIBFRAME Update, ALA Annual 2018

Nathan B. Putnam
Director, Metadata Quality, OCLC

OCLC

Agenda

- Goals of the analysis

- Process

- What we learned

- Future experimentation

Goals of the analysis

- Evaluate what type of BIBFRAME data OCLC could produce using the LC converter

- Review of the LC conversion stylesheet
  - Were modifications needed to support OCLC use cases?

Process

- 11 million records containing LCCNs from the OCLC Research copy of WordCat
  - OCLC Research WorldCat includes additional identifiers
  - All formats included in the set
- Converted selected records to MARCXML
- MARCXML sent through LC BIBFRAME 2.0 converter

What did we learn?

- Work IDs are very important
  - Useful at the beginning of the process
  - OCLC clustering adds them to OCLC Research WorldCat
- URIs are very important
  - Several blank nodes without them
  - Blank nodes = non-interoperable
- Converter contains BIBFRAME item descriptions
  - But we were only using bibliographic records
  - This left gaps in the data

OCLC Research updates to the converter

- Populate 758 field with Work/cluster IDs

- Modified to look for $0 and $1

- Preferred URIs for VIAF and FAST

Future experimentation

- Look at indexing possibilities/goals/requirements

- URI clean-up and population to existing MARC records to reduce the number of blank nodes

- Continue working with the Library of Congress
  - With multiple extensions to and variations of BIBFRAME emerging, OCLC is thinking broadly as we build capabilities that support the complex landscape