

MARC Content Designation and Utilization

Inquiry and Analysis



Examining MARC Records as Artifacts Reflecting Metadata Utilization Decisions

William E. Moen

<wemoen@unt.edu>

School of Library and Information Sciences
Texas Center for Digital Knowledge
University of North Texas

Research funded by a National Leadership Grant from the Institute for Museum and Library Services. Additional support provided by the University of North Texas School of Library and Information Sciences and the Texas Center for Digital Knowledge.





The MCDU Project

- **MARC Content Designation Utilization**
 - Provide empirical evidence of catalogers' use of MARC content designation
 - Identify commonly used elements of bibliographic records
 - Contribute to community discussion about core elements in MARC bibliographic records
 - Explore the evolution of MARC content designation
 - Develop research approach to understand the factors influencing levels of MARC content designation use



Metadata Record as Artifact

- Metadata creation as process
- Decisions by metadata creators
- Influenced by... ?
- Artifact reflects decisions, policies...
- Artifact reflects metadata utilization decisions
 - Decisions to use or not use available metadata elements



Metadata Utilization Analysis

- Developing a re-usable methodology
- Developing software tools and database design for storing large datasets
- Identifying questions to be answered
- Determining methods for analyzing data to address questions
- Compiling results



Questions to Answer

- Two categories of questions
 - General profile of the dataset
 - What is the distribution of records by Type of Record?
 - What is the distribution of records by Encoding Level?
 - Occurrences of content designation structures,
 - What is the number of total occurrences of all control and data fields and how many unique field tags are used?
 - In how many and in what percentage of records is each unique field/subfield combination used at least once?



Richness of MARC

MARC 21 Field Groups	Currently Defined	Obsolete	Total	MARC 1972 (Books Format Only)
00x	6	1	7	3
0xx	238	7	245	28
1xx	66	1	67	40
2xx	137	32	169	15
3xx	109	32	141	4
4xx	69	0	69	37
5xx	323	38	361	8
6xx	184	5	189	66
7xx	452	47	499	41
8xx	141	20	161	36
TOTAL	1725	183	1908	278



Occurrence Summary

Frequency	# of Fields/ Subfields	% of All Occurrences
> 600,000	1	4.4%
500,000 > 599,999	0	0%
400,000 > 499,999	13	39.9%
300,000 > 399,999	6	14.3%
200,000 > 299,999	6	10.6%
100,000 > 199,999	10	10.3%
TOTAL	36	79.5%

- Only **4%** of all fields/subfields account for **80%** of all occurrences
- **96%** of all fields/subfields account for only **20%** of all occurrences



Example Results

- 7,595,887 LC-created records in dataset
- Type of Record: Book, Pamphlets, and Printed Sheets
- Total number of unique fields: 167
- Number of fields accounting for 90% of occurrences: 14 fields (8.3%)
- Number of fields accounting for 80% of occurrences: 21 fields (12.6%)
- Approximately 110 fields occur in less than 1% of all records.

[Note: Fields are cataloger-supplied, not system-supplied]

Field Tag	Number of Records Where Each Field is Used at Least Once	Number of Total Occurrences of Each Field	Cumulative Total Percentage of Field Occurrences
650	5,387,282	11,778,732	10.910%
008	7,595,887	7,595,887	17.945%
245	7,595,887	7,595,887	24.981%
010	7,595,726	7,595,726	32.016%
300	7,586,264	7,586,415	39.043%
260	7,585,926	7,585,928	46.069%
050	7,027,027	7,095,639	52.642%
100	5,626,011	5,626,018	57.853%
500	3,264,297	4,582,571	62.097%
020	3,845,934	4,235,426	66.020%
082	4,034,888	4,036,101	69.758%
043	3,665,624	3,665,626	73.154%
504	3,373,297	3,403,714	76.306%
700	2,312,712	3,240,072	79.307%
880	512,563	2,327,504	81.463%



Implications

- Empirical basis for decisions about core elements in a metadata scheme
- Profiling repositories of metadata for aggregators
- Organizations can use methodologies and tools to analyze local utilization levels
- Contributions to changes in cataloging rules, practices, policies, and standards



References

- MARC Content Designation Utilization Project
 - <http://www.mcd�.unt.edu/>
- Assessing Metadata Utilization: An Analysis of MARC Content Designation Use
 - http://www.unt.edu/wmoen/publications/MARCPaper_Final2003pdf.pdf