

Understanding Support of FRBR's Four User Tasks in MARC-Encoded Bibliographic Records

by Shawne D. Miksa

Functional Requirements for Bibliographic Records

The MCDU (MARC Content Designation Utilization) Project [1], funded by a National Leadership Grant from the U.S. Institute of Museum and Library Services, is a systematic examination of MARC use through a quantitative analysis of over 56 million bibliographic records from OCLC's WorldCat database. The overarching research question for the project is "What is the extent of catalogers' use of the content designation available in *MARC 21 Format for Bibliographic Data* [2]?" One of the planned deliverables of the project is a list of frequently used MARC elements for bibliographic records representing 10 different material formats. A natural extension of this effort is to determine the actual usage of elements that can logically support the four different resource discovery tasks (Figure 1) defined in the *Functional Requirements for Bibliographic Records* (FRBR) [3]. Delsey [4] had already proposed a mapping of the MARC designations to the FRBR tasks as part of a broader study of the relationship between FRBR and MARC, and that mapping is available as a database from the Library of Congress's MARC site [5].

At the 9th Annual Conference of the International Society for Knowledge Organization (ISKO), July 2006 in Vienna, Austria, we presented data illustrating how catalogers' encoding of bibliographic data may or may

Shawne D. Miksa (smiksa<at>unt.edu) is assistant professor at the School of Library and Information Sciences, University of North Texas. She teaches and researches in the areas of information organization, specifically library cataloging and classification. She is currently researching catalogers' utilization of MARC 21 as well as other cataloging tools and resources.

not assist end users' tasks [6]. During the course of this analysis we realized that a fundamental question that needed to be asked was "What does it mean to support a user task and how many MARC data elements are needed in that support?" By pairing the MARC data elements

associated with the four user tasks with the frequency count data from our analysis, we are able to add another layer to Delsey's functional analysis of MARC 21.

While Delsey's analysis examines three categories of user tasks (resource discovery tasks, resource use tasks and data management tasks), our interests focus on the category of resource discovery. A complete description of the methodology is in the conference paper. Table 1 shows the total number of elements, gathered from Delsey's report, that support each of the four user tasks.

There are approximately 2,000 fields and subfields defined in MARC 21. We needed to minimize discontinuities between MARC categories and the parameters of our own format-specific sets. The project

team therefore chose to focus this part of the analysis by considering only the variable data fields and related subfields since the structure of these elements is common to all MARC material types as well as our 10 format-sorted record sets. Furthermore, in order to highlight the most frequently (commonly) occurring elements, we chose to set a threshold based on a minimum frequency of use of each field/subfield combination, eliminating from consideration the more rarely occurring combinations.

Using MARC records for books that were generated by OCLC member libraries (as opposed to the Library of Congress) as an example, we can see how actual cataloger >

FIGURE 1. The Four User Tasks in Resource Discovery ([4], p. 10)

Search: Search for a resource corresponding to stated criteria (i.e., to search either a single entity or a set of entities using an attribute or relationship of the entity as the search criteria).

Identify: Identify a resource (i.e., to confirm that the entity described or located corresponds to the entity sought or to distinguish between two or more entities with similar characteristics).

Select: Select a resource that is appropriate to the user's needs (i.e., to choose an entity that meets the user's requirements with respect to content, physical format, etc., or to reject an entity as being inappropriate to the user's needs).

Obtain: Access a resource either physically or electronically through an online connection to a remote computer and/or acquire a resource through purchase, license, loan, etc.

TABLE 1. Number of Elements Supporting User Tasks as of February 2006

MARC 21 Bibliographic Format	FRBR Resource Discovery Tasks			
	Find (Search)	Identify	Select	Obtain
Total no. of elements (fields, subfields, fixed field positions and indicator positions) that support a given task	454	972	375	468

from WorldCat, we have found [7] that catalogers' actual utilization of fields and subfields is not directly aligned with the fields and subfields prescribed in standards such as the Program for Cooperative Cataloging (PCC) BIBCO [8] and CONSER [8], and the National

and Minimal Level Record Requirements [10]. Hence, the question of user task support as it relates to field/subfield choices comes into sharp focus.

The entity-analysis model in FRBR "focuses upon the needs and interests of the user... provides a powerful framework to evaluate and define metadata content and structure provisional displays of bibliographic entities with their relationships...its emerging strength resides in how it defines bibliographic relationships and how they could be portrayed through powerful collocation displays of numerous expressions of different works." ([11], p. 14). The model, however, is at times necessarily unclear, or vague, due to the inevitable variations in information resources and their intended or actual use; FRBR does not, and cannot, provide an absolute true model of these complexities. For example, Jones [12] points out that "FRBR does not explicitly distinguish *expression* from *work*, noting that the conceptual boundary between these entities is culturally determined." (p. 227). In his article Jones focuses on the FRBR model as applied to continuing resources, and in his discussion he highlights that "the most important consequence of both the spare application of the principle of personal authorship to C[ontinuing] R[esources] and the exceptional concept of the *serial work* is that several user objectives of FRBR (and of the Paris Principles) are undermined, in particular, the user's ability to retrieve all the *expressions* of a *work*, to identify the *works* of a given author, and to select from among these the most appropriate to his or her needs." (p 235)

Catalogers' handling of such processes as encoding authorship is one facet that greatly affects a user's ability to find, identify, select and obtain resources based on authorship. Carlyle [13] points out that the incorporation of the user-defined tasks into the FRBR model assumes that "explicitly defined user tasks will facilitate use in catalogs that implement it." (p.272). MCDU has provided some evidence on how that implementation may or may not be happening within the context of cataloger utilization in a very large set of bibliographic records. At the very least, we have a beginning point from which to provide some meaning to the measurement of successful support of the four tasks. >

field/subfield usage breaks down by user task (Table 2). This particular subset of records accounts for 34.5 million of the 56 million records in OCLC's WorldCat database. For purposes of reporting we state that a task is "supported" by fields/subfields but without a clear definition of what that actually means.

TABLE 2. Number of variable field/subfields supporting resource discovery by task with number and percentage above the MCDU threshold

FRBR Task	Number of variable field/subfield data elements supporting it	Number (percent) that are commonly occurring (above the MCDU threshold)
Find (Search)	382	61 (16%)
Identify	683	82 (12%)
Select	172	28 (16.3)
Obtain	295	37 (11.9)

The findings raise the question of what these levels actually mean in the overall picture of cataloger utilization of MARC 21. For instance, we know what data elements are aligned with the user tasks, as per Delsey, but do we know how many of these field/subfield combinations are needed to support a user task. Are tasks with a higher percentage of combinations more strongly supported in practice than those with fewer? Can we say the range of 12%-16% signifies a high percentage or a low percentage of support? Furthermore, to what extent does this support (stronger or weaker) affect the possible expression in current MARC records of the entity-analysis model as outlined in FRBR – its four levels of entities (works, manifestations, items, expressions), their attributes and the relationships between the entities?

Cataloger decisions about what access points to include in a record are supposedly guided in large part by standards describing the mandatory fields and subfields that are required in any given catalog record for any given format. For example, in a MARC 21-based record, an encoding level (Leader/17 – fullness of a record) of "4" with an authentication code of "pcc" in the 042 \$a indicates a core level bibliographic record that is "...*authenticated under the auspices of the Program for Cooperative Cataloging*" ([2], www.loc.gov/marc/relators/relaothr.html#rela_042b). However, in the analysis of the entire 56 million records

Resources Mentioned in the Article

- [1] *MARC Designation Utilization Project* [website]. www.mcdu.unt.edu/?p=30
- [2] Library of Congress, Network Development and MARC Standards Office. (2006). *MARC 21 format for bibliographic data*. Washington D.C.: Library of Congress Cataloging Distribution Service. Retrieved June 22, 2007, from www.loc.gov/marc/bibliographic/ecbdhome.html
- [3] International Federation of Library Associations, IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998). *Functional requirements for bibliographic records: Final report*. Retrieved June 22, 2007, from www.ifla.org/VII/s13/frbr/frbr.pdf
- [4] Delsey, T. (2003). *Functional analysis of the MARC 21 bibliographic and holdings formats. Second revision*. Prepared for the Network Development and MARC Standards Office, Library of Congress. Retrieved June 22, 2007, from www.loc.gov/marc/marc-functional-analysis/functional-analysis.html
- [5] Library of Congress, Network Development and MARC Standards Office. (2006). Access 2000 database filename: FRBR_Web_Copy.mdb, updated 07 February 2006 [Data file]. Retrieved June 22, 2007, from www.loc.gov/marc/marc-functional-analysis/functional-analysis.html
- [6] Miksa, S., Moen, W., Snyder, G., et.al. (2006). Metadata assistance of the Functional Requirements for Bibliographic Records' four user tasks: A report on the MARC Content Designation Utilization (MCDU) Project. In G.Budin, C. Swertz, & K. Mitgutsch (Eds.). *Knowledge organization for a global learning society: Proceedings of the 9th International Conference for Knowledge Organization. International Society for Knowledge Organization 9th International Conference, Vienna, Austria. July 5-7, 2006* (pp. 41-49). (*Advances in Knowledge Organization, 10*). Würzburg: Ergon. Retrieved June 22, 2007, from www.mcdu.unt.edu/wp-content/ISKO06sdrm23May2006REVISED.pdf
- [7] Moen, W. & Miksa, S. (2006). *Commonly used MARC elements: Report. MARC Content Designation Utilization (MCDU) Project*. Retrieved June 22, 2007, from www.mcdu.unt.edu/?p=45
- [8] Program for Cooperative Cataloging (PCC), (2005). *Introduction to the Program for Cooperative Cataloging BIBCO core record standards*. Retrieved June 22, 2007, from www.loc.gov/catdir/pcc/bibco/coreintro.html
- [9] CONSER. (2005). *Record requirements for full, core and minimal level records*. Retrieved June 22, 2007, from www.loc.gov/acq/conser/recordreq.html
- [10] Library of Congress, (2006). *MARC 21 national level record and minimal level record requirements*. Retrieved June 22, 2007, from www.loc.gov/marc/bibliographic/nlr/
- [11] Madison, O.M.A. (2006). Utilizing the FRBR framework in designing user-focused digital content and access systems. *Library Resources and Technical Services, 50*(1), 10-15.
- [12] Jones, E. (2005). The FRBR model as applied to continuing resources. *Library Resources and Technical Services, 49*(4), 227-242.
- [13] Carlyle, A. (2006). Understanding FRBR as a conceptual model: FRBR and the bibliographic universe. *Library Resources and Technical Services, 50*(4), 264-273.