Notes on Operations
Can RDA Content, Media, and Carrier Coding Improve Discovery Facet Mapping?

Carolyn McCallum, Kevin Gilbertson, Steve Kelley, and Lauren E. Corbett

Online public catalogs have provided users with the option to conduct faceted searches for more than a decade. Although faceting is undoubtedly useful to the discovery process, the authors found that their system's default facet mapping was inadequate for their researchers' needs, particularly for the faceting of bibliographic formats, and librarians at their institution have worked extensively to revise this mapping. These revisions have relied on creating complex decision trees, which require the system to consult multiple fields and subfields in bibliographic records to assign more precise format facets. When their authority control vendor offered to add Resource Description and Access (RDA) coding to their bibliographic records, including the new Content, Media, and Carrier fields that describe formats with greater granularity than the General Material Designation, they questioned whether the new RDA coding might improve their public catalog's format faceting. They found that the limitations of the MARC format as a data encoding standard meant that the RDA coding was not appreciably more useful to the format faceting process.

The online public catalog interface of the Z. Smith Reynolds Library at Wake Forest University (WFU) has provided users with the option of faceted searching since 2009. Although faceting is undoubtedly useful to the discovery process, we found that our system's default facet mapping was inadequate for our researchers' needs, particularly regarding the faceting of bibliographic formats, and our librarians have worked extensively to revise this mapping. These revisions have relied on creating complex decision trees, which require the system to consult multiple fields and subfields in bibliographic records, to assign more precise format facets. When our authority control vendor, Backstage Library Works, offered to add Resource Description and Access (RDA) coding to our bibliographic records, including the new Content, Media, and Carrier Type (CMC) fields that describe formats with greater granularity than the General Material Designation (GMD), we questioned whether the new coding could be used to improve the format faceting in our public catalog. With this research question in mind, we sent our bibliographic records to Backstage for RDA enrichment.

Setting

Located in Winston-Salem, North Carolina, WFU is a private institution with approximately 4,800 undergraduate and 2,800 graduate students. Three libraries—a medical library, a law and professional library, and the Z. Smith Reynolds Library (ZSR)—support the university's academic activities. ZSR, the largest of the three libraries, serves both undergraduate and graduate students in WFU's

Carolyn McCallum (mccallcj@wfu.edu) is Cataloging Librarian for Nonprint Resources, Z. Smith Reynolds Library, Wake Forest University; **Kevin Gilbertson** (gilberkm@wfu.edu) is the Web Services Librarian, Z. Smith Reynolds Library, Wake Forest University; **Steve Kelley** (kelleys@wfu.edu) is Head of Continuing Resources and Database Management, Z. Smith Reynolds Library, Wake Forest University; and **Lauren E. Corbett** (corbetle@wfu.edu) is the Director of Resource Services, Z. Smith Reynolds Library, Wake Forest University, Winston-Salem, North Carolina.

College of Arts and Sciences, School of Business, Graduate School of Arts and Sciences, and Divinity School.

ZSR currently holds approximately 1.9 million print volumes and provides access to more than fifty thousand electronic journals (e-journals) and almost eight hundred thousand electronic books (e-books). Nonprint collections (film, microform, music, digital, etc.) and the university's archival and special collections (rare books and manuscripts) are also housed in ZSR. Additionally, ZSR has been a selective depository for US government documents since 1902. The library is organized into seven departments—Administration, Access Services, Digital Scholarship, Research and Instruction, Resource Services, Special Collections and Archives, and Technology—that regularly collaborate on library projects and initiatives, including the focus of this case study and analysis.

ZSR's integrated library system is Ex Libris' Voyager, and ZSR has used VuFind, an open source discovery system developed by Villanova University, since 2009 as its primary online catalog interface. With sophisticated indexing and versatile searching capabilities, VuFind enables ZSR librarians to customize the catalog experience via a number of algorithmic parameters, including variables in the SolrMarc software used to index MARC metadata. Moreover, VuFind provides progressive search refinements within sets of search results via multiple flexible query facets.

RDA for original cataloging was adopted at ZSR in December 2013, after accepting RDA for copy cataloging at an earlier date. These relatively small additions of RDA and RDA-hybrid records to our catalog meant that the large majority of our bibliographic records were fully Anglo-American Cataloguing Rules, 2nd ed. (AACR2)–compliant prior to the Backstage enrichment project in December 2014.

Literature Review

Only a few papers proved relevant to our specific goal, but others provided related ideas or problems, and were included in the literature review. We were interested in addressing feedback from librarians and the technical feasibility to quickly make any changes that we perceived as improvements in the granularity of facets, in particular to separate VHS from DVDs and music CDs from vinyl record albums (LPs). Our goal was not to conduct a usability study since we had recognized problems during our own searches in VuFind, nor was it to compare integrated library systems or to look ahead to linked data, but solely to examine whether it was possible to make practical and immediate changes with RDA-enhanced MARC records in VuFind to improve faceting.

Nelson and Turney explored the incorporation, use, and value of faceted navigation in the design of commercial websites. They observed three prominent characteristics in the sites' search interfaces: "(1) the importance of facets as a key component in the search design; (2) the personalization of the text that instructs the user; and (3) intelligibility of facet labels."[1] Applying their knowledge of e-commerce design and comparing it to the design of today's library discovery interfaces, the authors recommended three areas that both libraries and vendors must address and work together to improve: clarity of purpose and personalized instruction for the search box; selection and display of clear, meaningful, and jargon-free facet terms; and attracting users to the facets column to assist in narrowing or refining their search results.[2]

Hider approached the use of CMC fields in survey-based research that was designed "to map out catalog users' conceptualization of library resources, testing the content–carrier categorization proposed by RDA."[3] He concluded that content and carrier data combined does not come close to meeting searcher needs and that adding "additional facets, such as purpose and audience, would greatly enhance OPAC searching. Given their preponderance in this user group's ontology, they may in fact be as critical and as 'core' as the content and carrier facets."[4] Hider explained that "purpose" might be information versus entertainment and an example of "audience" was the visually impaired. He also stated that cataloging "rules do not prescribe the use of specific, standard taxonomies to express these facets, which is critical if the information is to be used in faceted navigation."[5]

Bernstein looked more generally at the limited utility of the CMC fields for meeting the researcher needs of finding, identifying, selecting, and obtaining materials, and argued for the increased use of the RDA carrier characteristics. He suggested that the MARC fields for the RDA carrier characteristics (340 – Physical Medium, 344 – Sound Characteristics, 345 – Projection Characteristics of Moving Image, 346 – Video Characteristics, and 347 – Digital File Characteristics) are not discussed much in the literature and barely used because, although the fields were "approved in July 2011 for inclusion in the [MARC] standard . . . they did not appear in OCLC's Bibliographic Formats and Standards (one of the primary references to which catalogers look when performing their work) until late July of 2013," and because of this "they have remained in the eyes of catalogers merely theoretical concepts."[6] Bernstein argued that the carrier characteristics supply needed detail for mediated materials to fully differentiate them and make them findable, over and above the more general level of detail in the CMCs, by providing "a necessary additional hierarchy level of description of, and access to, a resource's unique properties."[7]

Rice Sanders, working in Innovative Interfaces' (III) Encore discovery tool, briefly described one of the problems we were addressing (albeit as part of a larger improvement plan): having one umbrella label derived from MARC material types ("Web resource") instead of more granular terms

"such as e-book, e-map and e-journal. Now, with streaming video and other types of electronic content, the group [consortium] needs to agree upon labels for other kinds of electronic content."[8] Rice Sanders recognized that it would be necessary to make edits in III's Millennium to add new material types to introduce more granularity.

Belford offered a methodology to aid library professionals in the selection of a discovery tool. She discussed MARC Leader (LDR) and RDA elements, explaining that vendors may use different combinations of coding in their default facet mapping, and she offered samples for testing displays and results in systems. Belford noted that for music, medium of performance (in an optional MARC 048 field, or RDA MARC 382 field) and "MARC 344–347 fields (sound, moving image, video, and digital file characteristics)" could be useful in identifying formats if more of this data were present in records.[9]

Majors and Mantz looked specifically at discovery tools in searching for music, "where a keyword search will usually result in a multi-format set of results if not something richer and therefore more complicated. Empowering the user with effective tools to manipulate a large and varied search result set is key to user success with music searching."[10] Henry, also looking at music searching, but more specifically with regard to the effects of RDA, observed that the loss of GMDs from AACR2 removed the shortcut of adding "sound" to a search, which was counterbalanced by the ability to find the more specific format using facets. He explicitly stated that the CMCs were "not necessarily meant to be displayed in a public catalogue but instead could be used to generate more user-friendly descriptions such as 'compact disc.'"[11]

Ou and Saxon surveyed 1,300 III customers to learn how many chose to display CMCs in the public catalog. They called their survey results a snapshot. Out of fifty-three responses, thirty-three libraries (62 percent) reported that "they do not display the 336, 337, and 338 fields in their public interface at all."[12] Ou and Saxon noted that when a mixture of records—some with only GMDs, some with only CMCs, and others with hybridization (including both GMD and CMC)—exists in the catalog, it impacts public display. In survey comments, they received complaints about the workload related to coping with this mixture and of seeing "no appreciable benefit" from the changes.[13] They suggested that the "sustainable" option would be to add CMCs, noting that OCLC "anticipates removing GMDs from WorldCat records" sometime after March 31, 2016.[14] Additionally, they suggested that it might be possible to populate the CMCs in a systematic, automated fashion using a combination of fixed fields and other fields in the MARC record. They remarked that the CMC terminology, especially "unmediated," could be confusing to researchers. Format or material-type "icons" were generated from a single fixed field, that is, the "same way for both AACR2 and RDA records," and only one icon

could be generated per record.[15] Ou and Saxon offered that the CMCs might be an improvement in precision over the GMD, which provides either content or carrier, but not both, in a single display space, and suggested that "generating icons that are based, at least in part, on the Content, Media, and Carrier Types is a popular idea."[16] One survey response suggested that the "recently introduced field, the Form of Work stored in the MARC 380, as perhaps more useful than the Content, Media, and Carrier Types" because it "can include terms such as 'Play,' 'Television program' or 'Motion picture.'"[17] Ou and Saxon concluded that "this remains a time of transition" and that the "promise of the Content, Media, and Carrier Types and the FRBR entities they describe has not yet been fulfilled."[18]

Caudle and Schmitz discussed a shift to utilize the CMCs as the basis for format facets by writing new code to replace VuFind's indexing process, thereby simplifying the creation of the facets.[19] They worked to add the CMCs to AACR2 records via global edits that took more than a year to complete. They concluded that RDA improved format display but thought that they should do more to meet researchers' needs by improving the granularity of facets. Achieving this required the presence of CMCs in all bibliographic records and the development of additional complex coding. In pursuing these improvements, they found that a library's MARC record import script "will be just a little simpler," and a library "must decide if it is worth the amount of time and human resources necessary for implementation."[20]

Overall, the papers cited in the literature review matched much of our understanding of the problems to address, yet some voiced caution about the utility of CMCs in faceting. When Ou and Saxon suggested that it might be possible to populate the CMCs in a systematic, automated fashion using a combination of fixed fields and other fields in the MARC record, the authors were beginning an enrichment project with Backstage to do just that. Caudle and Schmitz delved deeply enough into facet mapping decisions based on CMCs to suggest practical and immediate changes that might improve the quality of faceting. However, their conclusions admitted that granularity remained problematic when using the CMCs exclusively to map facets.

Before RDA: Understanding Facet Mapping Options for Formats

After migrating to VuFind in 2009, the authors soon discovered that their initial facet mapping for books and films was not adequately granular to meet their researchers' needs and expectations. They created separate custom book search and film search boxes on the library website where researchers were funneled into selected channels, with pre-search facets

determined by the library. The range of materials included in such searches was not apparent to our librarians and researchers. For example, were monographic government documents included in or excluded from a book search? Were streaming videos included in a film search or excluded because they were online resources?

In fall 2011, prior to having RDA CMCs included in our catalog records, we reviewed VuFind's decision tree for MARC mapping. Specifically, we determined how to include streaming media in a film search and to separate e-books from other electronic resources (e-resources) (for example, journals, government documents, media). After reviewing VuFind's MARC mapping methods and MARC coding values, we added several refinements to better determine item format, using local cataloging practices and our desired outputs as guides. The determination largely relied on specific 007 code values (Category of material [subfield a] and Specific material designation [subfield b]) with a final inspection of the type of record (Type) and bibliographic level (BLvl) in the record leader.[21] These precise and accurate identifications in the back-end application established increased flexibility for managing granularity in displaying relevant and usable format facets in the user interface. Overall, the goal was to facilitate precision in searching and browsing ZSR's catalog. In 2012, we created a spreadsheet that highlighted the number of formats and the count of items associated with each format in the library collection and provided a basis for discussion of whether more granular format terms were needed to assist researchers in locating appropriate materials (see table 1).

The Question of "Format"

Our work in distinguishing bibliographic formats complemented the experience of many of the authors cited in our literature review. Customizing our catalog's faceting to create higher levels of granularity was a strong focus. While working on improving facet mapping, other questions became apparent: how do we define what is meant by "format"?; should we accommodate researchers' mental models, the "conceptualization" described by Hider, which might include factors such as audience; and how do we apply more than one format facet for a single record when desired?

Regarding format, depending upon an agency's or individual's use of the term, the meaning and definition can vary greatly. For example, AACR2's glossary defines "format" as "a particular physical presentation of an item."[22] OCLC's glossary defines it as "a standard for the representation and exchange of data in machine readable form."[23] In this paper, we primarily define format as the physical medium by which information is stored and presented, such as book, journal, microform, video recording, sound recording, map,

electronic resource, etc. These broad format terms can be further specified, for example: e-book, e-journal, streaming video or audio, microfilm, DVD, CD, atlas, and CD-ROM. As our work proceeded, we encountered cases where several factors, including researchers' conceptual models, determined how we presented an item's format in VuFind's facets, including sometimes assigning multiple format facets to a single record. We also recognized, as did Nelson and Turney, Hider, and Saxon and Ou, that the language used in facet labels should not be jargon heavy and difficult for researchers to understand. Furthermore, we knew that for certain resources, multiple facets would be applied, putting them in seemingly overlapping categories, such as being both a sound recording and an electronic resource for streaming audio.

Audiovisual Formats

As ZSR acquired a greater quantity of streaming videos, it became desirable to have those titles included in a film search. In VuFind's default mapping, all e-resource types—e-book, CD-ROM, database, and streaming video—were mapped to the electronic format facet. To identify streaming videos, we used coding from the 007 fields (subfields a and b) for video recording and e-resource, relying on the Specific Material Designation (SMD) to determine the class of video object. This clarity in format mapping was critical to our success in distinguishing various video recording formats, such as DVD, VHS, streaming video, and the generic video facet. Similarly, for audio formats, we used the 007 subfield d (for speed) to separate vinyl record albums (LPs) from audio CDs. Both LPs and CDs are mapped to the audio facet in addition to their separate facets for LPs and CDs. The ability to apply more than one facet to any single catalog record also aids the researcher in discovering a multiformat kit or a book with a supplemental CD-ROM.

Book with CD-ROM Supplement

In response to a problem reported by a research and instruction librarian, we reviewed the MARC mapping script and observed that a record for a book with a supplemental CD-ROM defaulted to the single facet "software" because it matched on the 007 coding values for "electronic resource." Further processing to determine additional facets was precluded because a facet value already existed. To account for individual catalog records that contain coding for more than one format, the MARC mapping logic was modified to allow for multiple facet assignments. In the case of a book with a CD-ROM, the modified methodology added a conditional check that pulled values from the 007 subfield a, along with the record's Leader values contained in the fixed fields Type and BLvl. This conditional allowed for and ensured more accurate identification of the mapping for a record's multiple

formats. Following these changes, a combination book and CD-ROM record mapped to both the facets software and book.

Government Documents

ZSR is a selective member of the Federal Depository Library Program and, like most libraries, uses the term "government documents" to describe publications of the US Government Publishing Office or by specific departments of the US government (for example, the Department of Labor), plus documents produced by any of the fifty state governments. As a special category or class of material, whose physical features vary depending on the format in which it is published, government documents themselves naturally are not addressed by the CMC fields. Using our definition of format as the physical medium, the term "government documents" would not have a separate VuFind facet. In the default mapping, government documents would be faceted by their physical formats such as e-resource, CD-ROM, microform, etc., according to the Leader or coding information in the 007, not according to who published these materials or their intellectual content. To support the research and instruction librarians' desire to separate government documents as an exclusive facet, the MARC fixed field GPub (008/28 Government Publication), was added into the MARC mapping to render government documents as an exclusive facet. For researchers using the VuFind interface to the library catalog, this meant that government documents would not appear in search queries refined with any other facet, such as book. This can be helpful when the quantity of government document bibliographic records is overwhelming in the search results.

We took an additional step to seek an even more precise way to map both print and electronic monographic government documents for the purposes of exclusion from the book facet. In addition to including the 008 GPub and 007 values for electronic resource in the MARC mapping decision tree, we added the 086 MARC field for Government Document Classification Number. This precision allowed us to exclude works created by the presses of state universities from our government documents facet to better fit ZSR's conceptual model of government documents. We discovered during the mapping process that some state university press publications were coded with an "s" in the 008 GPub denoting a state government document per OCLC's MARC Bibliographic Formats and Standards.[24] While not incorrectly coded as a state government document, the general perception among ZSR's librarians was that researchers would not recognize or regard state university press publications as state government documents. It may be argued that this situation arose because of our librarians' insistence on having a separate government documents facet, but the problem of potentially confusing our researchers remains without

Table 1. Abridged Formats and Item Counts Across Facet Mapping Revisions (as of 2012)

Mapping: Original	Mapping: Revision 1	Mapping: Revision 2
Book (861840)	Book (800321)	Book (800061)
Electronic (615320)	Electronic (2276)	Electronic (2519)
E-book (23267)	E-book (525635)	E-book (487633)
	Streaming video (2271)	Streaming video (2008)
		Government document (148880)
		E-journal (32591)

this accommodation. Overall, as seen in table 1, we felt we had improved the facets offered in VuFind, which would help save the researcher time, but we were not completely satisfied and wanted to explore the promise of RDA and the CMCs for further refining of our facets.

Introducing RDA Content, Media, and Carrier Type Fields into the Catalog

In early 2014, Backstage Library Works offered to perform a retrospective RDA enrichment of an entire catalog at no cost for current authority control customers. The enrichment would consist of adding RDA data elements to bibliographic records created according to AACR2 rules, thus making them RDA-hybrid records. Because the project would entail sending virtually all of the bibliographic records from our catalog to Backstage for processing, we decided to conduct the retrospective conversion in December 2014 after the end of WFU's fall semester to minimize any potential disruption in library services to our students and faculty. Before we sent our records, we first established a profile with Backstage detailing what changes we wanted to make to our records.

Completing our profile involved making dozens of decisions regarding the treatment of our records. One of the major decisions was to retain existing GMDs in the 245 subfield h. Although we could strip the GMD from records to make them RDA-compliant, we retained them because current catalogs present information in a manner to researchers that might cause confusion with the lack of the GMD. The other key elements of the enrichment processing specified in the profile included having Backstage convert 260 imprint fields to 264 imprint fields, spell out abbreviations and Latin phrases ("Dept." to "Department," "et al." to "and others," etc.), and add CMC fields.

It was the addition of CMC fields that led us to consider whether the inclusion of these RDA elements in our

bibliographic records would improve how VuFind performs faceting on our records. Our initial plan, developed before the RDA enrichment process was implemented, was to see how VuFind handled faceted searching in three scenarios: pre-RDA, post-RDA, and a combination. Pre-RDA would handle faceting the way we historically did; post-RDA would perform the faceting based solely on the CMC fields; and the combination would use both the pre-RDA and post-RDA methods.

However, we found value in adding CMC fields to our bibliographic records and wanted to sustain this practice. We were therefore pleased to realize that our ongoing quarterly authority control processing with Backstage included RDA enrichment of our bibliographic records at no additional cost.

Analysis of Vendor-Supplied Reports

In January 2015, Backstage returned more than two million processed bibliographic records and numerous reports to us. The 1,935 reports, with twenty-one different types, ranged from statistical analyses of the changes to a listing of all publisher imprint fields that were revised to a listing of all the physical description abbreviations that were spelled out (for example, "ill." to "illustrations"). For the purposes of this analysis, we considered the reports that indicated that a problem had occurred with assigning CMC fields to the bibliographic records.

The largest batch of relevant reports were those that indicated that CMC fields had not been added to a bibliographic record. A total of 356 records were listed on these reports. Of these records, 353 were for materials held in our Rare and Special Collections. These materials included papers, photographs, certificates, notebooks, and letters. We expected that these types of materials would be difficult for Backstage to parse and identify using their algorithms, particularly as the MARC 300 field physical description in the bibliographic records was either "folder" or "box(es)," not the more common physical descriptions such as "v." for volume or "disc." The remaining three items included two books and one DVD from the main collections. Of these, one book was partially cataloged, while the other book was part of a kit and inaccurately cataloged. The DVD was inaccurately cataloged, lacking both a GMD or 007 field, which made it difficult to identify as a DVD using an algorithm.

The next category of report was unrecognized GMD, meaning that the automated process failed to recognize the GMD included in the 245 field. Only thirteen records were included in this report, and none had CMC fields assigned. All thirteen records were from Rare and Special Collections and consisted of nine records with the GMD "Graphic," three records with "Microform Manuscript," and one with "Manuscript." These outdated or fabricated GMDs were added to records for locally held materials, with the belief that they would be limited to internal use within the WFU community. These codes were not intended to be processed by external computers and were not recognized by Backstage.

The final category of report for records that did not receive full processing was called "CMC Optional." Of the thirty-three records listed in this report, thirty-two were assigned a 338 field of "unspecified" and one record was assigned a 336 field that read "unspecified." We found that twenty-three of these records were for books and had a misapplied 007 field that should be applied only to media, three DVD records with the GMD in a foreign language, three records were for notated music and had incomplete 007 fields, three records were for pieces of equipment (the catalog is also used to track electronic equipment), and one record was for a US government document, which inexplicably had a German language GMD in the 245 field.

It became apparent that only bibliographic records that were already difficult or flawed had prevented Backstage from providing a thorough conversion to CMC fields. Problems such as an unusual physical description, an inaccurate or absent GMD, and/or an inaccurate or absent 007 field prevented the assignment of some or all of the CMC fields. What is remarkable is the low number of records involved. Only 402 of more tha two million bibliographic records were not assigned some or all of the CMC fields, or less than two-tenths of 1 percent of the records processed by Backstage.

Analysis of RDA-Enriched Bibliographic Records

In addition to analyzing the reports, we analyzed the changes made to our bibliographic records, focusing in particular on how the CMC fields were added. The bibliographic records were examined to determine if the correct CMC fields were added, corresponding to the format of the material described.

Based on random sampling, the majority of our records appear to have been processed with the correct CMC fields added. No mistakes were discovered in the assignment of CMC fields for books (print and electronic) and serials (print and electronic), the formats that constitute the vast majority of our collections. Sound recordings were processed correctly, as were video formats, including DVD, VHS, laserdisc, and streaming video. The only difficulty with the DVD format involved Blu-Ray discs, which must be coded by a cataloger as Blu-Ray in the 347 field. We discovered that the majority of our bibliographic records, and many OCLC WorldCat records, lack the 344-347 fields, confirming Bernstein's and Belford's observations.

Although most of the formats were accurately processed, two formats were problematic: kits and microfilm.

Because only twenty-two titles in our catalog have the GMD "kit," we examined all these titles. Each was uniformly assigned the same CMC fields, three-dimensional form (336), unmediated (337), object (338), regardless of the kit's actual content. Proper cataloging practice requires adding CMC fields for each type of item contained in the kit (booklet, DVD, CD, flash cards, etc.). None of these kits were simply three-dimensional objects. However, as we confirmed with Backstage, their system was only capable of adding one set of CMC fields per bibliographic record. Human intervention will be required to assign additional CMC fields for these kits. Any other multiformat materials (books with supplemental CD-ROMs or DVDs with extensive booklets included), even if they are not coded as kits, will also require human intervention to ensure proper assignment of CMC fields.

The microfilm format presented far greater problems. Prior to Backstage's processing, all of our approximately eighty thousand microfilm bibliographic records included the GMD "microform" in the 245 field. Each of the records was also coded "a" in the 008/23 Form of Item to indicate "Microfilm." Some, but not all, of the records had an 007 field, with the code "d" for "Specific Material Designation" to indicate "Microfilm reel." After Backstage's processing, we found that numerous records with an 007 field indicating "Microfilm reel" were assigned the Carrier Type "unspecified." We also discovered cases where the record lacked an 007, yet the record was assigned the correct Carrier Type "microfilm." We cannot understand why the inclusion of the 007 field (which should solidify the case for identifying an item as a microfilm) would generate the Carrier Type "unspecified." In the long run, this problem may not be terribly important at ZSR, because, at the time of this writing, a large-scale project is underway to weed and reduce our microform collections, both fiche and film.

Discussion

When we began to review our analysis of the RDA enrichment reports and the enriched records, we were struck by the fact that the CMC assignment proceeded so smoothly. The vast majority of bibliographic records had CMCs added to them, and of those, a tiny fraction were assigned an incorrect term. Interestingly, the records for formats and types of material that we initially found problematic when using the 007, 008, and other fields to determine faceting, were for the most part, not problematic when it came to the assignment of CMCs. For example, streaming video materials, for which the faceting had to be adjusted so that they would be included in the films facet, rather than the electronic facet, were all assigned the correct CMC values: two-dimensional moving image (336), computer (337), online resource (338).

After the success of assigning CMCs in our bibliographic records, we began to carefully think through the application of the CMCs to facets and questioned the value of running the three experimental catalog faceting scenarios discussed earlier. We noted that if Backstage was able to add the CMCs with relative ease and accuracy based on the metadata in our bibliographic records, the CMCs did not provide new information. Rather, the CMCs repackaged data that was already accounted for in our facet mapping. While this new packaging may prove easier to manipulate in future catalog systems and may simplify the transition of data from MARC to BIBFRAME (which does not have the complicated coding of the MARC Leader, 007, and 008), at present it does not add much, if any, value.

Although we had hoped at the outset that running the three experimental scenarios would reveal useful differences, as the project advanced, we realized that this was not the case. We recognized that the inadequacies of the CMC-only approach to faceting were related to using just the data regarding the physical characteristics of a bibliographic entity. As an example of an inadequacy, government documents would not be faceted according to local preferences by using just the CMC. Although the CMC fields were correctly added to government documents, these fields described only the materials' physical format. However, the very nature of government documents as a category of library materials is based on the fact that these materials are published by governments (federal or state). The CMC fields offer no information as to the provenance of a title. Clearly, the pre-RDA enrichment approach to faceting would be necessary to properly assign the government documents facet. Caudle and Schmitz noted that each library needs to decide for itself regarding the expense of developing new coding for facet mapping based on CMCs, and we decided to work within our existing structure.

Another difficulty with testing a CMC-only approach is that there is no distinction between serials and monographs when relying solely upon CMC coding. Both serials and monographs are coded with CMCs text (336), unmediated (337), and volume (338) for print materials and text (336), computer (337), and online resource (338) for electronic resources. The Leader field is required to distinguish a serial from a monograph. This substantiates the inability to rely upon CMCs alone in providing facets based on publication format.

From these considerations, we realized it was unnecessary to run a test of how faceting would work using the post-RDA approach because we knew that it would be inadequate in several key areas. Additionally, with our realization that there was no essential difference between the data contained in the CMC fields and the various fields consulted in our facet mapping (007, 008, GMD, etc.), and that the finer granularity of faceting in the pre-RDA approach

was provided by consulting more fields in the bibliographic record, we decided that it was unnecessary to run a test comparing how faceting works in the combination approach versus the pre-RDA approach. That is, both the pre-RDA and post-RDA approaches rely on translating data from the MARC record to create the facet: by consulting a table of 007, 008, and Leader values in the pre-RDA method, and by marrying the three CMC fields in the post-RDA method. Any approach to faceting that would be useful to researchers would require consulting multiple fields and subfields within the bibliographic record.

Conclusion

In the end, we found that CMCs alone do not provide for sufficiently robust faceting of public catalog searches. Although CMCs are more granular and specific than GMDs, our pre-RDA faceting has long relied on consulting the 007, 008, 086, and Leader fields during indexing to determine the proper format facet to display. These fields would have to be used even if the CMCs provided the initial basis for our mapping decisions. Rather than use the CMCs, it is easier to continue using our pre-RDA facet mapping because it is adequate to meet our needs, albeit cumbersome. We successfully improved our faceting in many ways, such as separating music CDs from LPs and moving streaming video from e-resources to the film facet, but it required hours of labor by a cataloger and a programmer to revise the mapping.

Even though we currently are not utilizing the CMCs for faceting, we believe the addition of the CMCs will ultimately prove to be beneficial. Because the CMCs unpack the dense metadata about physical format encoded in a number of fixed and variable fields, they make data eye-readable, easier for programmers to utilize, are generally more forward-facing, and potentially more useful in next generation library systems. During this transitional period in the bibliographic world, the more rigorous structure provided by the CMCs readies our data for the approaching linked data environment.

Another way to enhance the structure of bibliographic data is to follow Bernstein's advice for catalogers to increase the use of the 340, 344, 345, 346, and 347 (or 34X) fields to record carrier characteristics. Similar to the CMCs, the 34X fields parse data that was relatively hidden throughout the bibliographic record. Following Bernstein's recommendation, we have begun using the 347 field to record Blu-Ray carrier characteristics. This improves the structure and consistency of our data because prior to the creation of the 347, Blu-Ray data was recorded in the 007 fixed field and/or the 538 note, neither of which is easily searchable or indexed. Although the 34X fields and CMCs improve the structure

of the data for physical characteristics that determine facets, they are not designed to describe the intellectual content of bibliographic entities.

The increased use of the relatively new 38X MARC fields could address this deficiency. They include field 380 (Form of Work), 381 (Other Distinguishing Characteristics of Work or Expression), 382 (Medium of Performance), 383 (Numeric Designation of Musical Work), 384 (Key), 385 (Audience Characteristics), 386 (Creator/Contributor Characteristics), and 388 (Time Period of Creation).[25] Like the CMCs and the 34X fields, the 38X fields repackage data previously scattered throughout the MARC record. Unlike the CMCs and 34X fields that structure data about the physical characteristics of resources, the 38X fields structure data about the intellectual content of resources, which may prove useful in faceting.

The 380 field for Form of Work, for example, can be used to record whether a resource is a play, a television program, a choreographic work, etc. It could be enormously useful to researchers to have a facet displayed in the catalog to quickly distinguish records for the novel versions from the film versions for a given title, or the play versions from the opera versions. Also, the 382 field for Medium of Performance records the instrumental or vocal performance medium for a resource. This information, if displayed in a facet, could be quite useful for researchers looking for solo piano performance recordings of a particular piece of music or full orchestral scores with vocal parts. The 385 field for Audience Characteristics could be used to generate facets that would allow researchers to quickly identify resources that are geared toward certain ages (children, adolescents, adults), occupations (painters, cinematographers, librarians), or other demographic groups. The 388 field for Time Period of Creation provides information that could be displayed in a facet that would allow researchers to narrow their search results to contemporary primary sources about World War II or to present-day resources about seventeenth-century history. The other 38X fields also offer intriguing possibilities for assigning facets dealing with the intellectual content of bibliographic entities. We recommend exploration of the advantages offered by the 38X fields as a useful direction for additional research.

References

1. David Nelson and Linda Turney, "What's in a Word? Rethinking Facet Headings in a Discovery Service," *Information Technology & Libraries* 34, no. 2 (2015): 77, https://doi.org/10.6017/ital.v34i2.5629.
2. Ibid., 82–87.
3. Philip Hider, "A Comparison Between the RDA Taxonomies and End-User Categorizations of Content and Carrier,"

Cataloging & Classification Quarterly 47, no. 6 (2009): 548, https://doi.org/10.1080/01639370902929755.

4. Ibid., 558.

5. Ibid.

6. Steven Bernstein, "Beyond Content, Media, and Carrier: RDA Carrier Characteristics," *Cataloging & Classification Quarterly* 52, no. 5 (2014): 467, https://doi.org/10.1080/0163 9374.2014.900839.

7. Ibid, 484.

8. Martha Rice Sanders, Bob McQuillan, and Amy Carlson, "On Beyond E-Journals: Integrating E-Books, Streaming Video, and Digital Collections at the HELIN Library Consortium," *Serials Librarian* 62, no. 1–4 (2012): 193, https://doi.org/10.1 080/0361526X.2012.652920.

9. Rebecca Belford, "Evaluating Library Discovery Tools through a Music Lens," *Library Resources & Technical Services* 58, no. 1 (2014): 57.

10. Rice Majors and Stephen L. Mantz, "Moving to the Patron's Beat," *OCLC Systems & Services* 27, no. 4 (2011): 282, https://doi.org/10.1108/10650751111182588.

11. Stephen Henry, "RDA and Music Reference Services: What to Expect and What to Do Next," *Fontes Artis Musicae* 59, no. 3 (2012): 264.

12. Carol Ou and Sean Saxon, "Displaying Content, Media, and Carrier Types in the OPAC: Questions and Considerations," *Journal of Library Metadata* 14, no. 3–4 (2014): 245, https://doi.org/10.1080/19386389.2014.990846.

13. Ibid., 248

14. Ibid.

15. Ibid., 249.

16. Ibid., 251.

17. Ibid., 251–52.

18. Ibid., 252.

19. Dana M. Caudle and Cecilia Schmitz, "Keep it Simple: Using RDA's Content, Media, and Carrier Type Fields to Simplify Format Display Issues," *Journal of Library Metadata* 14, no. 3–4 (2014): 228, https://doi.org/10.1080/19386389.2014 .984572.

20. Ibid., 234.

21. OCLC, "Summary of MARC Leader and 008 Field," *Bibliographic Formats & Standards,* last modified July 28, 2014, http://www.oclc.org/bibformats/en/fixedfield/008summary .html.

22. Joint Steering Committee for Revision of AACR, *Anglo-American Cataloguing Rules,* 2nd ed. (Chicago: American Library Association, 1998), 618.

23. "OCLC glossary," OCLC, accessed October 28, 2015, http://www.oclc.org/support/documentation/glossary/oclc.en.html.

24. "GPub: Government Publication," OCLC, accessed October 28, 2015, http://www.oclc.org/bibformats/en/fixedfield /gpub.html.

25. "3XX—Physical Description, Etc. Fields—General Information," Library of Congress, accessed March 28, 2016, http://www.loc.gov/marc/bibliographic/bd3xx.html.