MARC Reborn: Migrating MARC Fixed Field Metadata into the Variable Fields

Steven Bernstein
Elihu Burritt Library, Central Connecticut State University, New Britain, Connecticut, USA

ABSTRACT
Despite calls over the past decade and a half for MARC to be replaced with an encoding standard that is more in keeping with current metadata practices, the current standard has evolved in such a way as to render many of the arguments by those who call for its demise moot. A further revision to the standard is proposed to address the one remaining problem with MARC so as to allow it to better serve the needs of information seekers for years to come.

ARTICLE HISTORY
Received January 2015
Revised July 2015
Accepted July 2015

KEYWORDS
MARC; encoding standards; migration

Introduction: Must MARC die?
Roy Tennant’s provocative call in 2002 for the death of MARC\(^1\) remains a seminal moment in the history of librarianship as a popular rallying cry for much needed change in our bibliographic systems. While oft-cited as a slogan by those who hunger for drastic change in library metadata, much of what Tennant proposed in his article and its follow-up, “MARC Exit Strategies,”\(^2\) was evolutionary rather than revolutionary. Tennant’s definition of MARC at the time conflated “several interrelated things … the MARC syntax, the MARC data elements, and the Anglo-American Cataloguing Rules (AACR2).”\(^3\) In the years following the publication of “MARC Must Die,” two of the three component parts of Tennant’s MARC conflation have indeed died. Although, due to the developmental nature of Tennant’s proposals, it is perhaps more accurate to say that these aspects have undergone a “reincarnation.”

The reincarnation of MARC syntax began a few months prior to Tennant’s publication of these columns when in June 2002 the Library of Congress introduced MARCXML as a syntactic alternative to MARC 21.\(^4\) As an XML schema, MARCXML allowed for the metadata content it encoded to escape the bonds of the library catalog so that it could be used and reused in a myriad of ways. MARCXML records could easily be transformed into other XML-based metadata

CONTACT
Steven Bernstein
bernsteinstj@ccsu.edu
Elihu Burritt Library, Central Connecticut State University, 1615 Stanley Street, New Britain, CT 06050, USA.

Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/01639374.2015.1075642.

Published with license by Taylor & Francis. © Steven Bernstein
formats such as Dublin Core or Metadata Object Description Schema (MODS), mashed up with external content, and styled for presentation in any way imaginable. While it might technically have been possible to do these things with MARC 21, the learning curve required for information professionals both inside and outside of the library community to make use of the ISO 2709 standard was highly prohibitive. So advantageous was MARCXML over MARC 21 that vendors of Integrated Library Systems (ILSs) quickly adopted it as the de-facto standard. Indeed, most next-generation ILSs introduced in the past decade have made use of MARCXML architecture. With the adoption of MARCXML, the library metadata encoding standard became simplified enough to join the broader information ecosystem while still maintaining its highly valuable complexity.

The reincarnation of the AACR2 also commenced in 2002 with the introduction of a strategic plan for that standard. This served as a catalyst for the development of Resource Description and Access (RDA). Using the International Federation of Library Associations and Institutions’ (IFLA) Functional Requirements for Bibliographic Records (FRBR) as a conceptual model, one of RDA’s goals was to separate elements into categories based on whether they recorded a resource’s attributes on the work or expression level or on the manifestation or item level. A hierarchical approach was also taken in recording a resource’s material type. RDA aimed to accommodate a digital age in which a resource’s content was no longer necessarily tied to its carrier. Thus, elements were introduced for recording an item’s content, media, and carrier. Included as one of its guiding principles, RDA also placed emphasis on recording the nature of the relationships between a resource and its various associated entities. After a decade of development, the Library of Congress implemented RDA on March 31, 2013.

The reincarnation of MARC data elements is the one aspect of Tennant’s conflation that has yet to be realized. Their longevity in their original body is understandable: relatively ancient as they are, they still get the job done efficiently. They provide for as many as 1,000 unique fields, each with as many as 36 unique sub-fields. Their use of three-digit field designators allows for the standard to be understood internationally as numbers are the universal language. So effective are the MARC data elements that—in his critiques of MARC—Tennant had but a single criticism to say about this particular piece of the puzzle: “While MARC offers at least some facility for dealing with multiple scripts—such as a book title in Chinese and translated or transliterated title—it handles them in such a way that it can be difficult for them to be properly processed by software.”

The birth of MARC and its major problems

That Tennant found only one fault in the MARC data elements is not to say that they do not have major shortcomings. Quite the contrary: MARC is a product of the compromises that had to be made as part of the shift from card catalogs to electronic catalogs. One of the major benefits of this shift was that—whereas card
catalogs allowed for only a select set of bibliographic elements to be indexed (author, title, subject, and classification)—electronic catalogs allowed for the possibility of providing an index for all properties that resources in a collection might possess, be they common or obscure. For this new advantage of electronic catalogs to be leveraged, however, the AACR2 content standard would have needed major additions and modifications so that those elements which were either not included or included as free-form notes on catalog cards could instead be recorded using controlled vocabularies in electronic bibliographic records. However, constrained as it was at the time by having to serve two masters (both traditional card catalogs and the electronic catalogs that would slowly but surely replace them), the Anglo-American Cataloguing Rules could not adapt as quickly as needed to provide this functionality. Inclusion of these elements would have been impractical given the tiny three-by-five inch size of catalog cards; indexing of these elements would have been impractical given the incalculable number of additional cabinets that would have been needed to store the even more incalculable number of additional access point cards in each of the additional indices.

Because these metadata elements could not be made available to users by way of the AACR2 content standard, they instead became a part of the developing standard that would be native to the electronic environment that made them possible in the first place: the MARC encoding standard. MARC came into being during an era when data storage and transmission lines were quite limited. As such, these metadata content elements were recorded using single-character codes in the standard’s fixed and 007 fields rather than in controlled terms that would be much more legible to humans in the standard’s variable fields. Though it was possible for these single-character codes to be represented in both the cataloging interface and the online public access catalog (OPAC) displays by textual labels, they—for the most part—remained apart from what the user saw when viewing a bibliographic record. The fixed and 007 fields were instead used by systems for advanced searching and report generation. If an element in the fixed or 007 field represented a property of a resource that AACR2 deemed important for the purposes of description, it was often also recorded textually in a free-form note field. Although necessary, this also represented a major drawback to the usage of MARC as the requirement to record the same property of a resource in two disparate fields—one as description to be understood by a human and a second time as an access point to be understood by a computer—greatly diminished the potential productivity of the cataloger. However, more important than how these elements were recorded or the impact their requirement had on cataloging efficiency was the fact that their inclusion as a part of MARC intermingled elements that were inherently metadata content into a format that was intended as a standard for metadata encoding.

While MARC’s dual role as both an encoding standard and a content standard has served an important function by allowing the library world to transition from card catalogs to electronic catalogs, this transition is now—with perhaps a few
exceptions—complete. The profession has moved to RDA, a standard that is “designed to take advantage of the efficiencies and flexibility in data capture, storage, retrieval, and display made possible with new database technologies.”

Unbound by the necessity to accommodate the maintenance of card catalogs, MARC—specifically the MARC data elements—can now be reformatted as a pure encoding standard and perhaps reincarnated as an Resource Description Framework (RDF)–based standard for the Semantic Web such as Bibliographic Framework (BIBFRAME). Before that can happen, however, it will be important to unify all of our metadata content under a single content standard so that the obsolete elements of the current iteration of MARC may be shed without losing important access points in the process. Many of the innovations of RDA were positioned with this very goal in mind. The question is: To what extent do those elements and values that are new to RDA match up to MARC’s fixed and 007 fields and values? Is there a one-to-one correspondence? If not, which elements and values need to be further enhanced to allow for the MARC elements’ reincarnation?

This article aims to answer these questions by assessing the prospects for batch migration of metadata content stored in the MARC fixed and 007 fields into the appropriate MARC variable fields containing RDA values. As this assessment is a comparison of two field/value pairs across a large series of particulars, it is best communicated in a tabular fashion. The conversion tables showing these parallels—one for each possible value of the 007 subfield “a” as well as one for the fixed field—can be found in the appendix to this article, in the supplemental material online. It is suggested that the reader review the tables at this point before proceeding with the main body of the article, which contains their analysis. Please note that the scope of this article is limited to the MARC elements for the bibliographic format. Migration scenarios pertinent to MARC authority, holdings, classification, and community formats are not covered. An assessment of the prospects for batch migration of the metadata content stored in the fixed-length data elements of these formats is left as a question for examination in the future.

**Analysis of the conversion tables**

Suffice it to say, there is no one-to-one correspondence between the metadata content stored in the MARC fixed and 007 fields and available MARC variable fields containing RDA values. While match points do exist on many field/value pairs, many do not have equivalents. In cases where MARC 007 or fixed field field/value pairs have no corresponding RDA element, the conversion tables note one of the following six reasons:

1. **No equivalent element:** RDA has no element corresponding to the MARC field in question

   *Example:* The MARC 007 field for motion pictures includes a subfield “f” for indicating whether the sound of the motion picture is on the medium or
separate from it. RDA does not include an element for recording this information.

2. **No equivalent value:** RDA has an element corresponding to the MARC field in question, but does not have a controlled term corresponding to the MARC code in question.

   *Example:* The MARC 007 field for maps includes a subfield “f” for indicating whether the cartographic item is a facsimile or other type of reproduction. This subfield corresponds very nicely to the RDA element for recording the Generation of the resource’s carrier (encoded in 340 ǂj). However, none of the terms used in the MARC Documentation to define the single character codes for the subfield match up to controlled vocabulary terms approved by RDA for use in this element.

3. **No equivalent field:** RDA has an element corresponding to the MARC field in question, but MARC does not have a variable field in which the RDA element can be encoded.

   *Example:* The MARC fixed field, “BLvl,” for recording a resource’s Bibliographic Level corresponds nicely with the RDA element for Mode of Issuance (RDA 2.13.1.3). There is, however, no place besides the fixed field “BLvl” for encoding this metadata.

4. **Split values:** The MARC code in question corresponds to two or more possible RDA terms.

   *Example:* The MARC 007 field for non-projected graphics includes a subfield “e” for indicating the primary support material for the image. This subfield corresponds very nicely to the RDA element for recording the Base Material of a resource’s carrier (encoded in 340 ǂa). Moreover, the terms used in the MARC Documentation to define the single character values for the subfield match up nearly perfectly with the controlled vocabulary terms approved by RDA for use in this element. That is, with one exception: a value of “c” (Cardboard/illustration board) corresponds to two separate terms in RDA (“Cardboard” and “illustration board”).

5. **Merged value:** Two or more MARC codes correspond to the one term.

   *Example:* The MARC fixed field, “LTxt,” for recording the form of a literary text expressed on a spoken sound recording corresponds very nicely to the MARC 655 subfield “a” for recording the form of a resource. Many of the terms used in the MARC Documentation to define the single character values for the field match up with specific Library of Congress Genre/Form Terms that may be used in this subfield. Two “LTxt” terms (Autobiography and Memoirs), however, both translate into the same genre term: Autobiographies.

6. **Administrative metadata:** The MARC code in question represents an attribute of the bibliographic record itself rather than the resource it records.
Example: The MARC fixed field, “Desc,” for recording the Descriptive Cataloging Form indicates the content standard used in creating the bibliographic record. It is inherently part of the encoding standard and does not need to be migrated to the content standard.

The solutions to the first two problems are simple and can be addressed with broad strokes by the Joint Steering Committee for Development of RDA (JSC). The JSC simply needs to update the standard to include these missing elements and values. The conversion tables included in the supplemental material point to cases where a term used in the MARC Documentation to define a single character code is a perfect candidate for inclusion in the RDA vocabularies with a note of “Proposed RDA Value.”

The solution to the third problem is also simple. The MARC Advisory Committee can resolve this problem by updating the standard to include these missing fields.

The solution to the fourth and fifth problems is a bit more complex as it requires knowledge of the individual resources that the effected records represent in order for a decision to be made with regard to migration. Transferring the metadata in these cases may need to be done on a record-by-record basis to avoid any loss of accuracy in the access points.

The sixth problem is not a problem at all. As such, no solution is needed to address it.

**Library of Congress—Program for Cooperative Cataloging Policy Statements**

In some cases, so as to allow for a better match between a MARC 007 subfield code definition and a controlled term, RDA instructions were disregarded in favor of related policy statements put forth by the Library of Congress—Program for Cooperative Cataloging (LC-PCC PSs). These cases are noted in the migration tables with a reference to the relevant policy statement [e.g., LC-PCC PS (7.13.2.3)].

**Mixed, multiple, not applicable, unspecified, unknown, and other**

Nearly every MARC 007 subfield permits codes indicating that the characteristic being recorded is either: (a) an amalgamation of more than one possible value (often a code of “m”); (b) not applicable to the resource in question (often a code of “n”); (c) unknown (often a code of “u”); or (d) of a quality other than those listed as possible values (often a code of “z”). In migrating this metadata, these values can either be retained (using as values in the variable subfields the term used in the MARC Documentation to define the code), or discarded. If this metadata is retained, the catalog display should be modified to provide context to it so that the user can readily identify what aspect of a resource is unknown, not applicable, and so on.
Manual entry

Indication that a resource is a festschrift is provided in RDA by the presence of the relationship designator “honouree” appended to the access point for the person to whom the resource is a tribute. Because there is no practical way for a computer to identify which of the access points representing the other persons associated with a work is the specific access point that represents the individual in whose honor the resource is a festschrift, migration of the metadata contained in the fixed field “Fest” must be done manually in cases where the field contains a value of “1.” In such cases, the relator term is entered into a MARC 700 subfield “e.”

If … then … else conditional statements

The value of a MARC 007 subfield “a” indicates the category of material to which the resource belongs. The definitions of all subsequent subfields and values in the 007 field depend on the value contained in the subfield “a.” Each conversion table must therefore be nested in a series of if … then … else statements in order for a batch migration of an entire collection to work properly. Alternatively, a collection may be divided into separate files based on the records’ category of material and the conversions can be done in stages using the appropriate conversion table for each file. If … then … else statements also come into play in migrating 007 fields for sound recordings, where values of “m” or “s” in the subfield “f” can each correspond to two different values in the parallel RDA element depending on the value of the subfield “b,” which precedes them. The various conditions and their results for the conversion of the 007 subfield “f” for sound recordings are noted in Table 1.

Concatenation

One of the biggest obstacles faced in transforming our legacy metadata so that it can become as functional as possible for humans and computers alike is the fact that we did not leverage our encoding standard to its fullest potential from the start. Certain MARC variable fields that could have provided a high degree of granularity (and therefore accessibility) through the usage of as many of the 36 alphanumeric subfields available to them as possible were instead made to encode metadata in a more note-like fashion, integrating disparate properties into single fields and subfields.

Table 1. If … then … else conditional statements for migrating 007 fields for sound recordings to the variable fields.

| For a | If b has a value of | Then the corresponding RDA element has a value of |
|-------|-------------------|-----------------|
| m     | c (cylinder)      | 344  | d fine                                      |
| m     | d (sound disc)    | 344  | d microgroove                              |
| m     | e (cylinder)      | 344  | d fine                                      |
| s     | c (cylinder)      | 344  | d standard                                  |
| s     | d (sound disc)    | 344  | d coarse groove                            |
| s     | e (cylinder)      | 344  | d standard                                  |
System access to the properties described in these fields was provided through codes in the fixed and 007 fields, thereby creating the very problem of description and access being split between the content and encoding standards that we are now aiming to resolve. The fact that these properties are described together yet accessed separately creates an additional hurdle when attempting to detach the function of access from the MARC encoding standard: namely concatenation.

The most common example of a MARC variable subfield that would require concatenation of the values transferred to it from multiple fixed and 007 fields is the 300 subfield “b.” This subfield is a junk drawer of content and carrier characteristics, recording the “other physical details” of a resource such as, “illustrative matter, coloration, playing speed, groove characteristics, presence and kind of sound, number of channels, motion picture presentation format, etc.” As is the case with the subfield “a” that precedes it (see “Specific Material Designation vs. Carrier Types” below), migrating metadata from the fixed and 007 fields into the 300 subfield “b” poses a plethora of problems, of which concatenation is only the first. The other major impediment to relocating this metadata to the 300 subfield “b” is the fact that the records that would be modified by such a process would likely already include this subfield, requiring the routine to either overwrite the contents of the subfield entirely or to transfer only those values not already recorded there. Both of these approaches have potential pitfalls. The ideal solution to these problems would be for the MARC Advisory Committee to update the MARC encoding standard to include additional subfields in the 300 field for recording each of the unrelated properties that are currently all recorded in the subfield “b.”

**Mathematical operations**

The MARC fixed field “Time” records the duration of a videorecording or motion picture and uses as a unit of measurement minutes formatted in three digits (mmm). The MARC 306 subfield “a,” to which the metadata in the “Time” field corresponds, however, records duration in terms of hours, minutes, and seconds formatted in six digits (hhmmss). As such, the metadata must go through a series of mathematical operations as part of the migration process. These operations are demonstrated in the following Excel formula in which the value of the “Time” field is contained in cell A1. Note that the actual formula used to perform the operations will likely vary considerably depending on the programming language used in the migration process.

\[
= ((A1/60 - \text{QUOTIENT}(A1, 60)) \times 6000) + (\text{QUOTIENT}(A1, 60) \times 10000)
\]

**Dates**

A trinity of other MARC fixed fields that concern themselves with a chronological property of a resource (albeit on a larger scale than does the “Time” field) and that
represent a huge complication for migration are the “DtSt,” “Date 1,” and “Date 2” fields. These fields contain machine-readable metadata about what dates are associated with a resource and how those dates relate to it. The “DtSt,” “Date 1,” and “Date 2” fields are closely related to the now deprecated 260 subfield “c,” which duplicates this metadata in a supposedly more human-readable fashion. Under AACR2, the way in which a particular date or pair of dates was associated with a resource was not explicitly stated (with the one exception being the date(s) associated with a resource’s manufacture). As such the 260 subfield “c” of an AACR2 record would contain generic information about a resource’s date of origin. RDA improved on AACR2’s handling of the Publication, Distribution, and so on area by separating this metadata out into discrete elements containing Production Statements, Publication Statements, Distribution Statements, and Manufacture Statements. The MARC encoding standard was updated in the spring of 2011 to accommodate these new elements with the introduction of the 264 field for recording the Production, Publication, Distribution, Manufacture, and Copyright Notice of a resource. The major difference between the 260 field and the 264 field that replaces it is that the 264 field identifies the type of date element it contains by way of its second indicator. Although the 264 subfield “c” allows for more granularity in recording the date(s) associated with a resource, it still presents these dates in a textual format that relies on human perception to effectively communicate the nuance of the metadata. Transferring the values of these machine-readable fixed fields to the variable fields meant to contain human-readable metadata unlocks a Pandora’s box of troubles. An alternative approach would be to make use of the MARC 046 field for Special Coded Dates. Not only does this field encode dates in a way that is both machine-readable and—with a proper user interface—human-readable, it does so with a high degree of nuance with regard to the type of date(s) encoded. Unfortunately, the encoding of the 046 field does not perfectly parallel the encoding of the “DtSt,” “Date 1,” and “Date 2” fields. Moreover, the 046 field is just another case of the MARC encoding standard performing a function of access that should be handled by the content standard. In order for this metadata to be migrated so that it may be retained in a post-MARC environment, RDA will need to be further developed so as to include an element or elements for recording dates associated with a resource in a coded manner.

**Specific material designation versus carrier type**

For the most part, the codes in the MARC 007 subfield “b” for indicating a resource’s Specific Material Designation map nicely to the 338 subfield “a” for recording a resource’s Carrier Type. There are a handful of exceptions, however.

**Maps and non-projected graphics**

Rather than recording what RDA would consider a resource’s Carrier Type, the subfield “b” of 007 fields for these categories of material contains a code indicating a
resource’s Unit. RDA defines a Unit as, “the physical or logical constituent of a resource” (RDA 3.4.1.1). Units are always preceded by a number designating the quantity of the Unit(s) that make up a resource and are commonly encoded in MARC in the 300 subfield “a.” Migrating this metadata into the subfield “a” would be problematic, however, as the AACR2 records that would be modified by such a process would likely already include a 300 subfield “a.” The choices in such a case would be to (1) overwrite the contents of the subfield entirely (ill-advised, as the extent of the units and any subunits would be lost), (2) use a string function of some sort to (hopefully) replace only those characters representing the units while leaving the characters representing the extent of the units intact (ill advised, as it could be overly complex with the potential for loss of metadata), or (3) migrate the metadata into a second subfield “a” (ill-advised, due to the potential user-unfriendliness of having duplicate metadata displayed in the record). The simpler solution is to migrate this metadata into the MARC 300 subfield “f” for encoding Type of Unit. Although generally used only for cataloging archival material, use of the 300 subfield “f” should be more widespread for other categories of material as its use isolates the extent of a unit or subunit (encoded in subfield “a”) from the type of unit or subunit (encoded in subfield “f”). The higher level of granularity that results from using the subfield “f” allows for more precise manipulation of the metadata in the parent field than does use of just the subfield “a” alone. Although the RDA terms that would be used in the 300 subfield “f” when migrating the metadata from the 007 subfield “b” for maps and non-projected graphics would be redundant with the terms already present in the 300 subfield “a,” having the terms repeated in this subfield would allow the terms to be more easily indexed, as they are not joined in a text string with numerical values and punctuation. Employment of the 300 subfield “f” for this purpose would likely require some tinkering with the record display on the local system so as to prevent the appearance of extraneous metadata. Suppressing the subfield “f” from displaying in the OPAC should prove to be a much less daunting task than suppressing a repeated subfield “a.” Caution should be taken, however, if using this solution for collections that are predominantly archival.

Globes

The subfield “b” of the 007 field for globes is interesting in that—rather than recording a characteristic of a resource’s carrier—the codes it contains indicate an aspect of a resource’s content: its subject matter. The codes specify the type of celestial body or bodies that the globe depicts: Celestial globe, Planetary or lunar globe, Terrestrial globe, or Earth moon globe. As these are merely subject headings, migrating them should—in theory—be a simple matter of including their equivalent values in the appropriate MARC 65X field. The problem for those who use the Library of Congress’ thesauri, however, is that the Library of Congress Authority Files do not treat globes in a consistent manner. Celestial globes are considered to be a resource’s form while Lunar globes are regarded as a resource’s subject subdivided by its form.
Terms for Planetary or lunar globes and for Terrestrial globes are non-existent. Complicating matters even further is the fact that the *Library of Congress Subject Headings* prohibits the use of the term “Globes” as a free-floating form subdivision for such resources, preferring instead the broader term “Maps.” For the sake of simplicity the recommendation put forward here is to migrate these values as the form of the resource into MARC 655 fields (per Table 2). It should be noted that genre/form headings corresponding to three out of the four values being migrated either do not yet exist or have been deprecated. These values should be marked as being uncontrolled through having a value of “4” in their second indicator.

**Tactile material**

Rather than recording what RDA would consider a resource’s Carrier Type, the subfield “b” of the 007 field for tactile materials contains a code indicating the Form of Notation used in recording a resource’s content. RDA defines Form of Notation as, “a set of characters and/or symbols used to express the content of a resource” (RDA 7.13.1.1). Form of Notation concerns itself with the communication contained in a resource and is therefore intimately connected with the script of the language of the resource. The two elements are encoded side-by-side in the MARC 546 field with the language encoded in the subfield “a” and the Form of Notation encoded in the subfield “b.” Although 5XX fields are intended to contain notes describing a resource and therefore typically contain uncontrolled free-text, RDA includes a brief list of prescribed terms for recording the form of tactile notation. Requiring controlled terms in this subfield allows it to act as an access point to a resource’s Form of Notation. Interestingly enough, the subfield “d” of the 007 for tactile materials, which is meant to contain metadata indicating the class of braille writing also migrates to the MARC 546 subfield “b.”

**Text**

The subfield “b” of the 007 field for text does not record what RDA would consider a resource’s Carrier Type, but neither does it record any other single RDA element. Rather, the codes used in this subfield seem to be all over the map of RDA elements, with two values indicating a resource’s Font Size, another indicating a resource’s Form of Notation, and a third indicating a resource’s Unit.

**Table 2.** Genre/form heading terms corresponding to specific material designation codes in the 007 subfield b for globes.

| MARC 007 | Definition | MARC 65X |
|----------|------------|----------|
| a (Celestial globe) | A model of the stars, sun, planets and other heavenly bodies. | 655 _0 |a Celestial globes. |
| b (Planetary or lunar globe) | A model of a planet (not the Earth) or of a moon (not the Earth’s). | 655 _4 |a Planetary or lunar globes. |
| c (Terrestrial globe) | A model of the Earth. | 655 _4 |a Terrestrial globes. |
| e (Earth moon globe) | A model of the Earth’s moon. | 655 _4 |a Lunar globes. |
• **Font size:** A value of “a” or “b” represents the Font Size of a resource being either Regular Print or Large Print, respectively. RDA defines Font Size as “the size of the type used to represent the characters and symbols in a resource” (RDA 3.13.1.1). As regular print is assumed unless specifically stated otherwise, a value of “a” is not migrated. A value of “b” is migrated into the MARC 340 subfield “n” with a value of “large print.”

• **Form of tactile notation:** A value of “c” represents the Form of Notation of a resource as braille. It is migrated into the MARC 546 subfield “b” with a value of “braille code.” The rationale for including a value representing a tactile Form of Notation in this subfield seems to be based on the subfield containing metadata that records the resource’s accessibility features for the visually impaired (i.e., Large Print and Braille Code).

• **Unit:** Of course, it very well could be that the 007 subfield “b” for textual materials is just a catch-all for metadata about a textual resource otherwise encoded elsewhere. Such would be the rationale for including a possible value of “d” in the subfield for indicating that the Unit Type of the resource is loose-leaf. A 007 subfield “b” for textual materials coded as “d” is migrated to a MARC 300 subfield “f” with a value of “volume (loose-leaf).” See the earlier section on Maps and Non-Projected Graphics for more information on use of this subfield.

In addition to the MARC 007 subfield “b” codes to which the above examples are exceptions there are also a few fixed fields that—when they contain certain codes—correspond to a resource’s Carrier Type. In some cases, these codes record metadata that is a replication of the metadata encoded in the record’s 007 subfield “b.” A code of “j” in the “File” field for recording the Type of Computer File, for example, specifies that the resource in question is an online resource, something that is also specified in the record’s 007 field. In other cases, these codes indicate that the resource in question belongs to a particular grouping of Carrier Types but does not specify the exact Carrier Type that the resource actually is. A code of “q” in the “Form” field for recording the Form of Item, for example, indicates that the resource is contained on a tangible computer carrier. Such a carrier could be one of many Carrier Types in RDA: computer card, computer chip cartridge, computer disc, computer disc cartridge, computer tape cartridge, computer tape cassette, or computer tape reel. Since migrating the metadata from these fixed field field/value pairs would result in either duplicate or split values in the 338 subfield “a” they are excluded from this discussion.

**Variant codes/values of a single fixed field or subfield mapping to different RDA elements**

While there is generally a one-to-one correspondence between the MARC fixed field/007 subfields and RDA elements, some fields/subfields map to two or more entirely different RDA elements depending on the codes or values they contain. In
addition, there are some cases where an individual fixed field code will correspond to two separate RDA elements either in conjunction with each other (requiring an “and” statement in the programming of the conversion table) or in disjunction of each other (requiring an “or” statement in the programming of the conversion table). These deviations are primarily due to RDA’s departure from the flat model of physical description under which MARC was developed and its adoption of the new hierarchical understanding of Content, Media, and Carrier. Cases where variant codes or values of a single fixed field or subfield map to different RDA elements include:

**Form of item**

The MARC fixed field “Form” records the Form of Item of a resource, which is essentially just another way of saying Specific Material Designation. As is the case with the 007 subfield “b” discussed in the “Specific Material Designation versus Carrier Type” section above—this metadata primarily corresponds to the Carrier Type element in RDA and should for the most part migrate to the MARC 338 subfield “a.” However, because any code in this field that would migrate to the 338 subfield “a” would be a duplicate of the values migrated to that field from the record’s 007 subfield “b,” this field is not migrated if it contains any of those codes. An additional reason why the field is not migrated if it contains certain codes is that some of these codes indicate that the resource in question belongs to a particular grouping of Carrier Types but does not specify the exact Carrier Type that the resource actually is. Migrating the metadata from this field in such circumstances would result in split values in the 338 subfield “a.” The three possible codes of this field that do not map to the 338 subfield “a” and which therefore do migrate are codes of: “d,” indicating that the Font Size of the resource is large print; “f,” indicating that the Form of Tactile Notation of the resource is braille code; and “s,” indicating that the resource’s Media Type is computer.

**Type of record**

The MARC fixed field “Type” specifies the Type of Record that a MARC record is. Although the “Type” field is technically administrative metadata, it can also serve as descriptive metadata that conveys the specific type of material the record represents. When viewed as descriptive metadata, the “Type” field corresponds very nicely to RDA’s Content Type element, encoded in MARC in the 336 subfield “a.” The fit is not a perfect one, however. A code of “g” in the “Type” field indicates that the resource is a projected medium. It maps to a value of “projected” in the Media Type element of RDA and is encoded in the subfield “a” of the 337, not the 336. A code of “k” in this field indicates that the resource is a two-dimensional non-projectable graphic. Because the code records both the Content Type (still image) and the Media Type (unmediated) of the resource it is migrated to the 336 subfield “a” in conjunction with the 337 subfield “a.”
**Type of computer file**

The MARC fixed field “File” records the Type of Computer File that a resource is. All but one of the possible codes for this field map to the RDA element for File Type, which is encoded in MARC in the 347 subfield “a.” However—since that one code is not migrated anyway because it corresponds to a Carrier Type value that is a duplicate value of metadata migrated from the 007 subfield “b”—practically speaking, all possible codes for this field map to the File Type element in the 347 subfield “a.” Two of the possible codes of the “File” field that migrate to the 347 subfield “a” represent Type of Computer Files that are actually genres of File Types, not File Types in and of themselves. They are: “e,” indicating that the Type of Computer File is Bibliographic Data and “g,” indicating that the Type of Computer File is a Game. So as to capture both the genre aspect and the File Type aspect of these Type of Computer Files they migrate to a 6XX field in conjunction with the 347 subfield “a” (see Table 3).

**Videorecording format versus video format and encoding format**

Thus far, all examples of this phenomenon have been found in the fixed field. The only example of variant codes of an individual 007 field mapping to two or more different elements in RDA occurs in the subfield “e” of the 007 field for videorecordings. While the majority of the possible codes for this subfield map to the RDA element for recording a resource’s video format (encoded in MARC in the 346 subfield “a”), the two codes that represent videorecording formats that are contained on 4 3/4” optical discs map to the RDA element for recording a resource’s encoding format (encoded in MARC in the 347 subfield “b”).

**Remote sensing images**

None of the contents of the 007 field for remote sensing images have an equivalent RDA element. Moreover, MARC does not provide variable fields other than the uncontrolled, free-text 500 “General Note” field in which this metadata can be encoded. Major revisions will be required to both standards in order to enable the migration of bibliographic records for resources of this category of material.

**Conclusion: Reincarnating the MARC elements**

With the broad adoption of RDA that commenced in the spring of 2013, two of the three component parts of Roy Tennant’s definition of MARC have been reborn

| Table 3. File types and form subdivisions/headings corresponding to “type of computer file” values. |
|----------------------------------------------------|
| Type of Computer File   | 347  | 6XX  | 655  |
|-------------------------|------|------|------|
| e (Bibliographic data)  | data file | N/A  | Bibliographies |
| g (Game)                | program file | Computer games | N/A  |
just as he recommended they should be. With RDA Day One behind us, it is now time for the library world to focus its efforts on revising the third component of Tennant’s MARC conflation: the MARC elements, specifically the fixed and 007 fields. Continued reliance on the MARC encoding standard as a solution for providing access to our metadata is no longer practical. It is time for the contents of these fields to be migrated to the MARC variable fields so that our encoding standard and our content standard can each serve their singular functions independent of each other. Not only will such a migration facilitate higher efficiencies in cataloging, but—should current homicidal punditry regarding MARC prove true—it will also enable the transition of our metadata to an alternative encoding standard such as BIBFRAME, an encoding standard that aims to give our bibliographic metadata a new life on the Semantic Web. For this to happen both the RDA content standard and the MARC encoding standard must be enhanced with new elements, values, fields, and subfields so that MARC—the MARC syntax, the MARC data elements, and the content standard that accompanies them—can enjoy a new life… because great metadata elements never truly die, they just get placed in new fields.

Questions for further study: Other MARC formats

This article has explored the prospects for migration of metadata contained in the fixed and 007 fields of MARC bibliographic records into MARC variable fields using RDA values. It is likely that such a migration will also be necessary for MARC authority records, MARC holdings records, and records of other MARC formats. If such a migration is necessary, what are its potential benefits and drawbacks? What hurdles might be encountered in such a process?

Acknowledgments

The author thanks his colleagues and friends Richard Churchill, Lisa Plavin, Dana Tonkonow, and Nicolas Tomaiuolo for their review of this article prior to its submission.

Notes

1. Roy Tennant, “MARC Must Die,” Library Journal 127, no. 17 (2002): 26–28.
2. Roy Tennant, “MARC Exit Strategies,” Library Journal 127, no. 19 (2002): 27–28.
3. Tennant, “MARC Must Die,” 26.
4. “An XML Schema to Represent MARC Records,” Open Archives Initiative, last modified June 14, 2002, http://www.openarchives.org/OAI/2.0/guidelines-oai_marc.htm
5. “Strategic Plan for RDA, 2005–2009,” Joint Steering Committee for Development of RDA, last modified July 1, 2009, http://www.rda-jsc.org/archivedsite/stratplan.html
6. Tennant, “MARC Must Die,” 26.
7. “Key Features,” Resource Description and Access, last modified April 14, 2015, http://access.rdatoolkit.org/
8. The table for the fixed field refers to the individual fields using the labels found in OCLC’s “Bibliographic Formats and Standards,” https://www.oclc.org/bibformats/en.html
9. In some cases, the conversion tables will indicate the existence of a relevant controlled term external to the RDA standard. Such cases are designated with a note consisting of the term’s type followed by the specific thesaurus or heading system from which the term has been taken. For the sake of simplicity, the conversion tables include only those subject, genre, or name terms that emanate from the *Library of Congress Subject Headings*, the *Library of Congress Genre/Form Terms*, and the *Library of Congress Name Authority File*. Language terms are incorporated from the MARC Code List for Languages. Libraries that do not employ these particular thesauri will likely need to adapt the tables to their specific usage scenario.

10. “MARC 21 Format for Bibliographic Data: 300—Physical Description (R),” Library of Congress, Network Development and MARC Standards Office last modified September 2013, http://www.loc.gov/marc/bibliographic/bd300.html