Backstage RDA Enrichment Service

Last Updated: **March 4, 2020**

Since Day One on March 31, 2013, Backstage has enriched approximately **194,962,082 records** through our RDA Enrichment services, representing **218** of our clients and their data. The smallest file so far was only **1,000** records; the largest file was over **8.0 million**.

Academic libraries (46%) have, by and large, comprised the greatest share of records that have gone through processing. **Public libraries** (32%), on the other hand, have generally represented the set with the fewest participating members. **Special libraries** (22%) make up the remainder. Please note that special libraries, in this sense, are understood to be libraries that are not purely academic or public; for example, corporate libraries, consortia, law libraries, etc.

The purpose of this document is to provide our clients with an updated overview of the options that our **218** clients have chosen from our available online RDA profile. In addition, the profile also references our extensive wiki and guide, both of which contain more details, examples, and other information available about our processing.

The online RDA profile is comprised of six sections: Data Information, Validation Options, GMD to CMC, Descriptive Fields, Access Fields, and **Reports**. We will only concern ourselves with the first **five sections**. Our intent with presenting this information is to give our clients a better sense of what their colleagues have already chosen in regards to RDA enrichment. It can be helpful to see your own path if the way ahead has already been travelled.

All of our processing and profile options can be customized to your specifications. In fact, some of our options were borne from working with our clients on customizations they wanted included for their own files. And so we carried those over into the online profile for others to also take advantage of, as desired.

Database Information

File Delivery

Clients have the option to upload their records either through our online website or via FTP. The vast majority (72%) are choosing the **website**. Similarly, we are delivering most of our files back through the **website** (74%). The website can be convenient for most clients as the interface is intuitive and there is no time limit to when files are available for retrieval.

We have found that, typically, nearly every record will be enriched with RDA elements during processing. This is especially true if we are converting 260 fields to 264 fields and adding 33X Content, Media, and Carrier Type fields to the records. Clients have the option to receive all of their processed bibs back or just the changed records instead. More clients are choosing to receive all of their bibs (68%) rather than changed only (26%).

While we offer clients the option of adding a stamp to their processed or changed records in either the 040 field or 005 field, we have also had inquiries about adding stamps in 9XX fields. Clients that choose the **040 $d UtOrBLW** stamp (74%) are also more likely to also choose the **005 field update** (61%). However, a respectable number are also choosing to add a **9XX RDA ENRICHED** (29%), **9XX MARS** (16%), or adding a stamp to some other field (6%).

Please also note that we do not offer adding **040 $e rda** to records that are enriched via machine process. PCC recommends against this policy so it is not something that we draw attention to within the profile settings. Of course, we can still make this change, though we strongly encourage our clients to re-consider the use of another field instead.

File Format
Although our clients can send us their files in MARC-XML format, and some do, most send their records in MARC21 format. We can accept either as there is a one-to-one relationship between MARC-XML and MARC21; essentially, there is no loss of data when converting between the two formats.

The primary question we are also interested in is what *encoding set* those records will be both sent by our clients and delivered back to our clients: UTF-8 or MARC-8. More of our clients (37%) are sending their files in **UTF-8**, which was not the case 10 years ago. But times have changed and so have Library Management Systems (LMS); UTF-8 is better equipped to handle larger character sets and more data formats (e.g., XML, TEI, EAD, etc) are natively stored in UTF-8 format.

Incidentally, more clients want their files returned in **UTF-8** (38%) format. We can handle files that may be a mixture of both MARC-8 and UTF-8. In fact, some clients might send in one format but expect a different format upon delivery, which our system can handle without any diacritic issues being introduced.

Enrichment & Validation
Backstage offers options to tackle existing RDA records (040 $e rda) validation & error-checking as well as enriching non-RDA records with RDA elements. We are living in a world which largely consists of RDA records and hybridized non-RDA records.

Approximately 90% of our clients are choosing to enrich their (pre) AACR2 records with RDA elements. On top of that, 88% also choose to validate & upgrade select elements within *existing* RDA bibliographic records.

Validation Options
Updates & Corrections
Backstage can make changes in over 100 different MARC fields within your bibliographic records. These changes are made *prior* to any RDA enrichments that you may also be interested in applying. We have broken up these different sets of validation, updates, and punctuation changes into logical steps in our profiles. Of course, our clients can pick & choose from amongst and within these various options to let us know specifically what they want to see happen with their processing.

When it comes to validation corrections for numeric fields such as 010, 020, 022, and 034 fields, 82% are in favor; 82% choose to update leader & fixed fields; 83% update & delete various fields; 83% update & delete various subfields; 83% update 1st & 2nd indicators for various fields; 81% make special field conversions & additions; and 78% check indicators based on initial articles.

Overwhelmingly, our clients are choosing to do the vast majority of updates available in this option.

GMD to CMC
GMD Standardization
Backstage makes use of two lists for standardizing GMDs in 245 $h subfields: AACR2 and Common

Terms. The AACR2 List contains 33 standard items and the Common Terms list contains 75 different items. **Audiocassette** is acceptable under Common Terms, but would be changed to **sound recording** under the AACR2 list. Clients can also choose to supplement either list with custom terms. GMDs typically have brackets surrounding the terms and so there are a couple options to add or remove these as desired.

Most of our clients are choosing to use the AACR2 list (65%) when controlling their GMDS, though we do have some clients also choosing the Common Terms list (6%). Surprisingly, more (11%) are interested in having us apply custom terms for their GMD processing. About 67% of our clients want their GMDs to be enclosed within brackets, while 3% ask us to remove brackets surrounding GMDs.

GMD Processing
PCC recommended retaining the GMD within the bibliographic record until March 31, 2016, which was precisely three years after Day One. This should have given ILS vendors ample time to make use of the 33X fields for display or discovery purposes.

Most of our clients (61%) prefer to retain their GMDs, for now. The tide appears to already be changing with clients now considering removing the 245 $h from their records. Still, there are some clients (31%) that are asking us to remove their GMDs.

As part of this instruction, we can also *move* the existing GMDs down to a 500 note or 9XX field so they still reside in the record as a failsafe; 58% pick a 500 field, and 55% pick a 9XX field instead.

Aside from retaining, removing, or moving GMDs, we can also *add* a select number of missing GMDs to their records. Perhaps not too surprisingly, few clients (22%) want us to actually perform that service. For those few clients still interested, however, an even smaller group (3%) specifies that added GMDs should be applied to *AACR2 records only*.

CMC Processing
Our authorities team has put forth a lot of research and development into fully flushing out this particular option. Currently, there are over **1,400** separate rules that our system uses in order to find the most accurate representation of Content Type, Media Type, and Carrier Type terms to add to the record.

Nearly all of our clients (93%) embrace the idea of adding both **$a terms** and **$b codes** to 336, 337, and 338 fields. Backstage processing does not currently add *multiple sets* of 33X fields to the records it enriches. We have looked into doing this in a reliable and confident manner, but have been unsuccessful so far. Currently, we consider GMDs, fixed field values and 300 $a information.

Descriptive Fields
Title - 245 Field
Just because we have a lot of options available for title updates does not necessarily mean that our clients are fully invested in making all of those updates. About 63% are choosing all available updates. This means that the remaining 37% of our clients are breaking out the title options with more granularity.

Of those remaining clients, only 3% ask us to remove certain terms from 245 $b if the Leader byte 06 is not **e** or the 007 position 00 value is not **m**; 19% choose to have a new 246 field added when the 245 contains an equal sign (=) preceding the 245 $b; 9% want certain noun phrases removed from 245 $b and moved into $c instead; 27% adopt the practice of changing **[et al.]** to **[and others]**; 16%

prefer replacing **[pseud.]** with **pseudonym**.

Edition & Cartographic - 25X Fields

This option updates abbreviations commonly used in the edition statement (250) and cartographic mathematical data (255) fields. Although PCC does not currently recommend updating 250 fields via machine process, we still offer this option for clients interested in pursuing this action.

Abbreviations spelled out in 250 fields include numeric designations (e.g., 1st to first, etc), ed. to edition, rev. to revised , etc. In fact, our processing includes spelling out abbreviations for other language versions of certain terms in 15 total languages. Abbreviations spelled out in 255 fields include RA to Right ascension, Decl. to Declination, eq. to equinox, as well as others.

Approximately 71% of our clients choose to include **both** of these 25X field updates for their processing. That leaves us with 29% who consider only one of the fields instead. Of this latter group, 4% want updates made in 250 fields only, and 11% want updates in 255 fields only.

Imprints - 260 Field

Far and away, imprint fields are one of our most involved options for clients. There is much that can be accomplished or addressed in this one option. There is so much here that even addressing the **conversion** of 260 fields to 264 fields is handled separately from this option.

So far, over half of our clients (66%) select to have **all** of the following enrichments applied to their records during processing. That still leaves us with 34% who want to take things one step at a time with this option. Let's take a look at the results below of what these other clients are deciding to do.

Due to a change in ISBD, each element taken from outside the resource should be in its own set of brackets. So a 260 field with **[**New York, 1973**]** would be updated to **[**New York**], [**1973**]**. Out of the other 34% remaining clients, 62% of those are interested in applying this update.

Some of the older forms of Latin abbreviations, such as **S.l., s.n., n.d.** are now being spelled out since those older forms are not necessarily apparent to other users. For the clients that are electing to take matters into their own hands with this option, over half (66%) are making this update.

Sometimes an **i.e.** note is included in 260 $a. A minority group of 41% wants this information moved from the 260 field into a new 500 field instead.

RDA uses standard abbreviations for most places such as states, countries, etc. Place names that don't have a standard RDA abbreviation can be spelled out in the 260 $a & $e. Also, abbreviations such as **Dept.** and **Co.** that may be found in 260 $b & $f can be spelled out. For the $a & $e changes, 41% want this, while 50% want the updates for $b & $f.

There are many changes that can be made in the way dates are recorded from AACR2 to RDA in 260 $c: 54% choose to spell out abbreviations in the names of months of the year; 64% prefer to update **ca.** to **circa**; 52% want dates with "printing" to be used as a publishing date and enclosed within brackets; 64% approve of changing dates that use a dash to represent an unknown number within a date to a "between" date instead; and 61% instruct us to update copyright dates such as **c1984** to **©1984**.

LC Policy is to try to supply a date of publication if at all possible rather than use the phrase **date of publication not identified**. So, out of this remaining group of 34% that address the imprint options

one at a time, 63% want our system to update from **©2001** to **[2001], ©2001** and *not* add the phraseology mentioned above. This is a clear indication that, for this option, most of our clients are adhering to the LC policy on the matter.

Conversion of 260 to 264

After addressing all of the update options available for imprints, this particular question is rather straightforward comparatively: 73% ask us to convert their 260 fields to 264 fields.

Our profile lists the conversion as our default setting mainly because we want our clients to see what the conversion will look like during a test run, and then give them an option to decide for their full set of records after that point. However, PCC Policy currently recommends that 260 fields should **not** be converted to 264 fields via machine process.

Physical Description - 300 Field

A number of abbreviation updates can be applied, by subfield, to the 300 physical description field. There are **32** abbreviations that are spelled out in the 300 $a (extent) and 90% of our clients have asked us to make these changes, along with 90% interested in **28** updates to 300 $b (other physical details), and the 89% who want the **15** changes for 300 $c (dimensions).

There are also a handful of changes that can be made for Specific Material Designations (SMDs); **6** updates that change **sound disk** to **audio disc**, etc. So far, 31% of our clients are exploring this latter option.

Other Field Additions

As we continue to expand our RDA enrichment offering, we include fields that may not be as widely accepted but still garner interest from our clients. This next section details the interest in those efforts to this point.

The 046 field is useful when your dates cannot be correctly coded in X00 $d due to B.C. dates or incorrect dates being used; 24% have requested this update for their processing. Our system will create an 046 field from the 100 $d.

Similarly, we can look at other fields within the bibliographic records to add corresponding RDA enrichment that may be useful in a machine environment. Namely, we can add 382, 383, and 384 fields derived from data found in the 240 field.

About 30% of our clients want us to map the 240 **$m** to **382** to denote the *Medium of Performance*; another 28% ask us to map **$n** to **383** to denote the *Numeric Designation of Musical Work*; while 29% are in favor of mapping **$r** to **384** to denote the *Key* and help differentiate between identical musical works.

Note Fields - 5XX

For this particular option, there are **13** abbreviations that can be spelled out in 500 or 504 notes fields. Our team has taken great care to ensure that our system does not mistakenly expand an abbreviation that was actually part of a person's name within the notes field; this kind of issue is more likely to appear in 500 fields as the 504 abbreviations are more uniformly applied.

Around 23% of our clients are in favor of spelling out abbreviations in **both** the 500 and 504 fields. However, we also have the option to treat these two fields separately. Of the remaining 77% of our clients, 15% choose to update 500 fields only while 28% choose to update just the 504 fields.

RDA rules allow for the content of 502 fields to be broken up into separate subfields for specificity; 22% of our clients ask us to add subfields into their existing 502 fields.

Access Points

Conference Place Names

In AACR2, multiple place names in an X11 field were separated by a **comma** and the word **and**. In RDA, a **semi-colon** is to be used instead. Most of our clients (87%) are choosing to make these updates in **all** of the X11 fields.

Date Abbreviations

There are many different abbreviations in $d (i.e., date) of 1XX, 6XX, 7XX, and 8XX fields that can be spelled out. The majority of these date changes only take place within X00 fields, though there are month names that can appear in X10, X11, and X30 fields. When it comes to making these particular kinds of date updates, 89% are in favor of **all** changes.

Another aspect to this option is changing abbreviations *preceding* the dates in $d. For example, 30% ask us to update **b. 1945** to **1945-**; 29% want us to change **d. 1945** to **-1945**; and 30% choose to fix **fl. 1945** to **active 1945**.

Dept to Department

One of the most talked about changes in abbreviations has been the change from abbreviating the term **department**.

Although this change was made during the conversion of the NAF to conform to RDA, since the authorized access point is based on the *preferred name* which is based on *usage*, it is still possible for there to be headings that use the abbreviation. The spelled out form should be found in a 4XX field (within the authority), so changed bibliographic headings should be able to be matched and flipped to the proper form.

Nearly 88% of our clients are in favor of expanding **Dept.** to **Department**.

Uniform Title Updates

Certain abbreviations that appear in arranged statements for music can now be updated, such as acc., arr., unacc., and violoncello.

In addition to these, our system can change **$a Selections** to **$a Works. $k Selections** in 240 fields and **$t Selections** to **$t Works. $k Selections** in select fields. Our default is to make **all** of the above changes and 83% agree with that recommendation.

Coupled with this option is the ability to also convert any uniform title with two languages into *separate* 700 fields with their own respective $l (lowercase L) subfields. This is a relatively new addition to our service and 17% have taken us up on the offer to this point.

Relator Terms

In AACR2, many relator terms (aka relationship designators) were not included on headings. With RDA, the relator terms are **retained** and **expanded** if they exist in their abbreviated form. We have a list of **29** abbreviated relator terms that our system consults and updates your headings as necessary; for instance: **jt. auth.** being changed to **author**.

The majority of our clients (90%) want us to retain and spell out their abbreviated relator terms. Included in this instruction is a list derived from the Open Metadata Registry that our system uses to check for consistency with existing relator terms. The list contains over **460** unique terms; those that do not appear on the list are reported out.

Going one step further, our system can also convert $e relator terms to $4 relator codes or vice versa; 1% prefer the $e to $4 change while 48% elect to head the other direction with converting $4 codes to $e terms. Moreover, we can include both $e and $4 in the headings; while 5% have chosen this path, we do want to note PCC does not recommend retaining **both** $e and $4 in the headings.

Sacred Works

O.T. and N.T. should no longer be used, but the terms should be **spelled out** as *Old Testament* and *New Testament*, along with *Koran* changed to *Qu'ran*.

With AACR2, a Biblical book title within an X30 field had an additional subfield containing either Old Testament or New Testament. With RDA, the "Testament" subfield is removed when a Biblical book is in the following subfield. For instance, **$a Bible.** *$p O.T.* **$p Exodus, I-IX** is updated to **$a Bible. $p Exodus, I-IX**.

Our default processing includes all of the above updates and 91% of our clients follow this.

Update X00 $c

Finally, there are a number of new RDA rules relating to X00 $c. Our system can update this subfield to RDA standards, such as adding parentheses or capitalizing the terms listed therein. As Phase 3 of the NAF cleanup gears up for processing later this year (2015), part of the focus is on doing further work with X00 $c. Until that happens, we still have 86% interested in receiving this X00 $c update for their bibliographic records.

Conclusion

We remain committed at Backstage to providing the best possible RDA enrichment service for your library's needs. This document has been created in such a way that each time a new client has processed their full catalog through our enrichment service we will update these findings to reflect the new results.

As mentioned at the start of this document, including these details might help provide a **narrative** for the way in which most libraries choose to handle their own RDA enrichment with Backstage. And this kind of information might also assist you with your own decisions about what to pursue at this time.

Please contact us with any questions you may have: **800-288-1265**.

Sincerely
The Backstage Authority Control Team