

**To:** Joint Steering Committee

**From:** Jennifer Bowen, ALA representative

**Subject:** Content and Carrier Terms in RDA (Response to the GMD/SMD Working Group report, *5JSC/Chair/6/Chair follow-up*)

ALA thanks the Working Group for its thoughtful work on this complex and difficult topic. The Working Group's latest report provides us with a significant step forward, especially concerning the basic concepts of content and carrier and in increasing our flexibility to address combinations of formats. However, ALA is unsure that the report has fully succeeded in addressing our concerns. In particular, we have questions concerning how the elements will be integrated into RDA. ALA members are also concerned about how display and implementation issues will be resolved.

One of the most cogent points the report makes occurs in section 6.2,

Further exploration is needed to understand how to record and display "specific carrier" terms effectively.

Effective implementation of all the interactions between bibliographic level, remote/tangible access, content, and carrier remains a challenge. The difficulty lies in trying to force resources into a linear arrangement of distinct boxes/categories. What we are dealing with is actually a multi-dimensional matrix of such categories – and the dimensions keep increasing! The multiple meanings of digital – strictly digital content vs. traditionally analog content being stored digitally (and possibly created digitally) vs. digital and tangible or digital and remote distinctions – further complicate the situation.

## General Comments

While ALA understands that the Working Group was specifically charged with using the categories of content and carrier identified and defined in *5JSC/Restricted/Chair/1* as a starting point rather than undertaking further deconstruction of terms, some ALA members find that the report and its recommendations suffer from a lack of precision, rigor, and clarity that would have been greatly improved by adhering more closely to the FRBR model. Some are concerned that the terms selected, especially those related to carrier, do not actually represent carriers.

It is not totally clear in *5JSC/Chair/6/Chair follow-up* how content and carrier terms are going to be used within RDA. Currently, "form of carrier" is listed as a mandatory element under RDA 1.4, but "form of content" is not. (Note that the element is labeled as 'form of carrier' in RDA 1.4., not "type of carrier", as given at 3.3, or "type and form of carrier" as in the Working Group report). If this is deliberate, what is the rationale behind making one mandatory but not the other? The Working Group report is worded in terms

of “requiring” various elements related to content and carrier, and we are unsure whether this is a recommendation to add these elements to the list of mandatory elements in RDA 1.4, or simply a recommendation to have the guidelines for each data element worded in such a way that they instruct that the data element be recorded. If it is the latter, we recommend going one step further, to have the guidelines recommend that as many content and carrier terms be added to fully describe multi-type resources, with an option to record only one, as we will indicate below in our more specific comments.

ALA reviewers of the Working Group report had further difficulty comparing the report with the draft of part 1 of RDA because the terminology of the report and the terminology of the draft are not the same:

- Broad content term (report) vs. Type and Form of Content (RDA 4.2)
- Specific content term (report) – no data element in the draft of part 1
- Broad carrier term (report) vs. Media Category (RDA 3.2)
- Specific carrier term (report) vs. Type of Carrier (RDA 3.3)

It was sometimes unclear to reviewers where the various lists would be covered within RDA as a result.

ALA would prefer to see the terms “type and form of content” and “type and form of carrier” that are used in the RDA Prospectus be changed to “type of resource” and “form of carrier” respectively. Using the words “type” and “form” exclusively in the different categories will help to distinguish them from each other.

The appendices seem to have varying levels of specificity and it is confusing to have the same term used in so many different ways (e.g., the term *book*). Because these are closed lists, there is a need to develop a mechanism to add terms to represent emerging technologies either in content or carrier categories.

ALA is concerned that for multi-type resources, some of the recommendations in this report may be a step backward unless RDA includes additional guidance to help the cataloger sort through the many possibilities that currently are given in no order of precedence. The exclusions in the definitions are also problematic. ALA questions why we couldn't have content types *object* and *cartographic* used together, or *digital* and *sounds* or *digital* and *cartographic*. Restricting terms in Appendices A and B does not allow, for example, a cataloger to record that an audio resource is digital.

## **Section 1. Summary of Recommendations**

**1.1.** As we mentioned above, in RDA 1.4, only “form of carrier” is listed as a mandatory element for the things under discussion in this document. If the Working Group recommendation is accepted, it would then seem appropriate to include “type of resource” (our recommendation for the name of the content element) as a mandatory element as well. Whatever the decision regarding mandatory elements, however, ALA believes the instruction within RDA should be to record as many broad content terms as

apply to the resource, with an option to apply only one term.

There are too many terms for sound categories in the list of broad content terms; we would prefer to see the term *audio* here as a broad content term. In conjunction with this recommendation, we recommend moving the terms *spoken word*, *music recording* and *sounds* (rather than sound) and consider them “specific content types”. Note: the example in 2.1 where a user may want a music resource but can only use a Braille or an audio format illustrates our point that *audio* is a better broad content term. Trying to make a distinction between the terms *sound* and *audio* is problematic. *Sound* should be in the glossary as a reference to *audio*.

**1.2.** ALA has concerns that without a more cohesive approach to recording this data, any given type of data may not be recorded consistently. It may be that an element should be designed to cover the specific content type. Where and how that element displays is a separate issue.

**1.3.** Again, ALA recommends that the instruction within RDA be to record as many broad carrier terms as apply to the resource, with an option to apply only one term or to use the term *mixed material* (or another similar term; see our comments on Appendix A below).

While the terms listed here may correspond to terms in AACR2, many are no longer logical or comprehensible to the average user. Particularly problematic are the terms *audio*, *digital*, and *projected*. The status of *analog* needs to be discussed.

For specific comments on terminology, see ALA’s comments below on Appendix B.

**1.5.** ALA recommends that this instruction should not be an option, but, rather, the rule. The instruction should be to repeat content and carrier type elements as needed. The option should be to supply only one or to use the term *mixed material*.

**1.6.** ALA found it difficult to evaluate this recommendation, as we are unsure where in RDA it would be covered. We were also whether the recommendation, by omission, was suggesting **not** to display the broad carrier list at 1.3? We can see the value in not displaying these terms since they might be confusing to users, and may be more fit for filtering than for display.

## **Section 2. Justification for Requiring Content and Carrier Information Contents**

Concerning the third paragraph, ALA observes that the presence of the terms themselves doesn’t necessarily enable interoperability, unless there is a deliberate effort made to select terms that can be mapped to those terms used by other sectors and domains.

### **Section 3. Selection of Terms**

**3.1.** ALA finds that many of the recommended terms do not achieve the goal of intelligibility to users. Some of these terms will not be intelligible to even educated users who may be interested in such resources, while some terms will be interpreted by many users in ways different from our intended meaning.

At the same time, it is less important to find terms now that will be good for display than it is to design the data structures well. If the past is any clue, our data will likely last much longer than whatever presentation style we use now.

**3.3.** While we acknowledge that this was not part of the Working Group's Terms of Reference, some members of ALA would prefer to see the terms for content further deconstructed into separate lists representing work and expression.

**3.4.** ALA strongly recommends that any term lists within RDA should be actively maintained, and that RDA developers monitor for new and ceased term lists to refer out to.

### **Section 4. Use of Multiple Terms**

ALA agrees that these data elements need to be repeatable.

**4.1.** We agree that catalogers need to balance their effort with cost. We have our doubts, however, that "the likelihood that these terms will be individually indexed" or "whether the entered term would be useful to correctly identify the resource" should be a criterion. Considering whether a library's OPAC of the moment currently indexes a particular element is a rather short-sighted approach; we need to be thinking longer term. And it's not whether the term itself is useful that is important; the display can be changed as terminological trends come and go. What is important is that we think that internal or external users will find the concept/category useful, regardless of the term used to name it.

**4.2.** We question the necessity of defining the content terms to be artificially mutually exclusive. The real world and the resources within it are not mutually exclusive, and we need to accurately represent that real world.

This section shows another area where we need some mechanism for showing the relationships between data elements, as described in the ALA response to the draft of part 1 of RDA.

We find it confusing that the report uses terms as illustration here that do not appear in either of the appendices (e.g., *dual disc* on p. 7).

**4.3.** ALA strongly recommends that information about originals and reproductions should be recorded as separate elements, and identified as such.

## Section 5. Display Considerations

**5.1.** It is not clear what is being recommended to be displayed versus what is needed for search filters.

**5.2.** It is not clear from the examples how exactly the “second level” carrier types are to be recorded or displayed, or to which RDA rule they apply. For example, the movie on videodisc does suggest where DVD, NTSC, or region codes would be recorded or might be displayed. The terms in the audio tape example do not follow the form of any of the specific terms under audio.

## Section 6. Issues

**6.1.** ALA agrees that the term “content” is not a good choice. Most ALA members prefer the term “type of resource,” as mentioned above. This is similar to the Dublin Core element “type”, and we note that six of the elements in 1.1 are the same or very similar to Dublin Core:

- RDA data = DC Dataset
- RDA moving image = DC MovingImage
- RDA object = DC PhysicalObject
- RDA software = DC Software
- RDA sound = DC Sound
- RDA textual = DC Text.

However, some members find that such a broad name for a bibliographic data element may render it basically meaningless – and so would prefer at least to rename it “type of work/expression” or “type of content”.

**6.5.** ALA agrees that RDA development needs to take into account basic issues of encoding, a primary reason being the need to indicate the relationships between elements.

While the points made here are valid, the report fails to take a stance on whether the element should be a separate element or not. The guidance in RDA needs to be definitive and specific; otherwise, the sharing of records will be impaired, if not impossible.

## Appendix A

It is not clear what the highest level term would be for a computer or video game. *Mixed?* *Software?* The definition for *Software* seems to exclude games and digital multimedia, but the category of mixed seems too narrow.

**Comments on specific terms:**

**Audio:** The term audio is not a “carrier term”; it belongs in Appendix A, not Appendix B. We prefer this term as the broad term for the various sound terms in 1.1 (i.e., consider *music recording*, *spoken word*, and *sounds* specific content terms).

**Choreographic:** Do not restrict to notation; include also images.

**Data:** The term *data* is problematic in that any intellectual work could be considered *data*; we see the need for a concept for “raw data” or “native data” to cover factual information in a structured form. As defined here, we see the term *data* as a difficult concept for catalogers. Do not exclude cartographic resources.

**Mixed:** As we mentioned above, we recommend flipping the rule and the option around. The default should be to record as many terms as apply, while the option should be using *mixed*. We recommend that the term be changed to *multiple* or *various* or *mixed material*; mixed has a connotation of inseparably mixed into one, like a website, which does not work as well with collections of various objects, which themselves fall into one. We suggest changing in the definition, “Examples include archival resources containing a variety of materials, ...”, and “instructional kits, computer and video games, digital multimedia, and websites with moving image, sound, and textual information.”

In the category under “excludes”, in the option, the example is given of *textual – spoken word* to describe a book and read-along CD packaged together. The form of notation given in this example (i.e., the dash between two instances of the same element) conflicts with the notation used to denote a hierarchy of elements. In this example, *spoken word* is not a more specific kind of *textual*, it is a different kind of content. Separating them with commas would be preferable (*textual, spoken word*) because they are equivalent terms indicating broad “type of resource”. The hyphen as used in other examples seems to indicate a hierarchy, whereas commas indicate equivalent terms.

**Moving image:** Not all moving images consist of “a series of visual representations”. Some digital moving images rather are algorithms that instruct a computer to change different parts of the screen at certain times.

**Music notation:** In the “Content-Carrier Examples” there is a hyphen after *music notation – tactile – Braille*. Is this example complete?

**Music recording:** We recommend that this be considered a “specific content” term and that the broad content term be *audio*.

**Object:** We recommend removing mention of dimensionality from naturally-occurring; some of this type of object (such as a leaf) can be as two-dimensional as a painting. Naturally-occurring objects do not have to be solid; they can be liquid or

gaseous. Do not exclude cartographic resources.

**Sound:** recommend replacing this term with *sounds*.

**Spoken word:** Delete. Use *sounds + textual* (or *audio + textual*). We recommend that the definition include “recordings of lectures and presentations”. Is this meant to cover moving images of sign language communication? If so, the name needs to be changed. If not, which category would such resources be in?

**Textual:** Change to *language* or *linguistic*. It is not clear what the purpose of “primarily verbal” is. What needs to be excluded? Latin is now primarily a written language rather than a verbal one, but we still want to call it textual.

**Visual:** Recommend replacing this term with *still image*. Do not exclude cartographic resources.

## Appendix B

ALA agrees with the recommendation to refer out to controlled lists of terms. Some terms in the specific list have a statement that allows the cataloger to refer to another thesaurus to identify a useful term; all terms should include such a statement.

However, we find that this set of terms and definitions needs significant work. Many of the terms do not describe carriers, either broad or specific. Some different terms denote the same carrier. A better title for Appendix B. might be “Closed List of Broad and Specific Terms for ‘Form of Carrier’” (instead of “... ‘Type and Form of Carrier’”).

We note that some of these terms will not be intelligible to even educated users who may be interested in such resources

- 1 aperture card
- 2 cartographic
- 3 DAT
- 4 learning pack
- 5 microopaque
- 6 photomechanical print
- 7 projected
- 8 radiograph
- 9 remote-sensing image
- 10 SACD
- 11 wire recording

while some terms will be interpreted by many users in ways different from our intended meaning.

- 1 audio
- 2 digital
- 3 graphic

- 4 icon
- 5 kit
- 6 manuscript
- 7 paper roll
- 8 part
- 9 section
- 10 view
- 11 visual (if we continue to exclude moving images from this category)

### Comments on specific terms:

**Audio:** *Audio* describes a type of expression, not a carrier; it belongs in Appendix A, not Appendix B. We prefer this term as the broad term for the various sound terms in 1.1 (i.e., consider *music recording* and *spoken word* specific content terms. The term *audio* in normal English implies a sound resource, and thus would not seem to indicate either the inclusion or the exclusion of aspects such as analog vs. digital recording, audio CD vs. MP3 file, etc. Essentially, *audio* is being used to refer to all of the “mainstream” types of commercially available physical sound media, and *digital* to refer to the less typical types. (Where would MP3 CDs fall in this scheme?) This division is not easy for catalogers to grasp, and it seems even less likely that users will be able to understand it.

**Cartographic:** This division has one flaw that is easily corrected. Cartographic resources have been divided somewhat arbitrarily. It is logical for *globe*, *model*, and *raised relief model* to be included in the broad term *three-dimensional*, but dividing the remaining specific terms between *printed* and *graphic* is not at all logical. The vast majority of all cartographic resources are “produced through a printing or other mechanical process” (*printed*), including the specific carrier terms listed under other broad carrier terms, except of course *manuscripts* which is handled under that broad term, and *remote-sensing images*. By definition, cartographic resources portray graphic representations combining geographic information and map-making techniques. The logical home for all the cartographic specific terms, except those under *three-dimensional* is *graphic*. Moving the terms *atlas* and *map* to the broad term *graphic* and deleting the reference to “cartographic images” in the definition of *printed* will resolve this problem.

**Digital:** To use the term *digital* in a way as to exclude audio CDs does not seem logical, since by their very nature audio CDs are digital media. Do not exclude *audio* or *projected* resources. Whenever something is defined as *digital*, it will need to have an additional format statement like *online* or *CD-ROM* – otherwise it will be ambiguous as to whether it is direct or remote access.

The specific terms under digital also have many problems:

What is the intended distinction between *disk* and *diskette* and *magnetic disk*? *File* and *website* are not carriers, at least using the AACR2 definition. It is not clear why some file types include the word *file*, some include the word *format*, and some

include neither What does *optical disk* refer to, given the many separate terms for CDs and DVDs? *Computer game* is not a carrier. Perhaps what is meant is *computer game cartridge* or *computer game disc*? Perhaps the Nintendo Gameboy and Xbox need to go under something else like these headings. While some game discs are clearly described as *CD-ROMs* or *DVD-ROMs*, some, such as XBOX, are not so identified and are difficult to identify as anything other than “disc” and the platform name. Our understanding is that some sets of these terms (such as those for DVDs) are the same carrier, used in different ways. If *e-book* and *e-doc* are used, *e-serial* and/or *e-journal* should be added.

**Graphic:** *Graphic* describes a type of expression, not a carrier. These terms do not denote types of carrier: *album*, *art original*, *art reproduction*, *chart*, *diagram*, *icon*, *portfolio*, *profile*, *section*, *view*. These terms describe type of carrier along with non-carrier information: *activity card*, *art print*, *flash card*, *picture*, *remote-sensing image*, *sketchbook*, *study print*, *wall chart*.

**Manuscript:** This is an arbitrary definition that is not helpful in describing resources. We do not see the logic behind including texts, maps, and musical scores, but excluding other types of resources. The medium for most modern manuscripts (paper) is very rarely made by hand. *Printout* and *typescript* denote the same carrier – paper. The vast majority of non-manuscript printed resources nowadays are made by typing or keying. Is a published book that has handwriting instead of type to be considered a manuscript? If a poet handwrites one of her previously published poems, is that published or unpublished? (In other words, is publication an attribute of a work, and expression, or a manifestation?) For a document consisting of 23 pages of type on letter-size paper, stapled in the upper left corner, how is the cataloger supposed to know whether it is published or unpublished, whether it is printed from a computer or photocopied or printed by a printing company, whether the document is “personal use”? Eliminate this term, or recast it much more narrowly to early manuscripts. Replace it with concepts such as type (*handwriting* vs. *machine type* (optionally broken down further)) and publication status (*published* vs. *nonpublished*)

**Microform:** Some members feel that *Microform* does not denote a carrier, but rather an aspect that nearly any carrier can have. A *microopaque*'s carrier is paper, while *microfilm*'s is film. Regarding the example: *Book* is not a type of *microform*. What is the definition of *book* here? Is it a work- or expression-level term, or a manifestation-level term? It appears that information about originals and reproductions is being mixed up, which should be recorded separately, and identified as such.

**Multimedia:** Do not exclude *computer game* resources. *Board game* describes a type of carrier along with non-carrier information. *Kit* does not denote a carrier, but the fact that a resource is made up of multiple pieces that are used together in some way. Similarly, *Learning pack* does not denote a carrier, but purpose of the work.

**Printed:** Under specific terms in the 3rd column, these terms do not denote types of carrier: *piano* [*violin, etc.*] *conductor part*. These terms describe type of carrier along

with non-carrier information: *atlas, large print, map, photocopy, score* (and its child terms). It is not clear what is meant by *part* here. It is not clear what the difference between *score* and *sheet music* is. It would be helpful to add *serial* and two categories of *serial, periodical* and *newspaper*, as examples.

**Projected:** The definition of *projected* does not even mention the concept of projection. It uses the term *display*, which is really the only thing that all the items dumped in there have in common in terms of use. The label itself is completely artificial as is the category as well. *Projected* describes a manner of use of a manifestation, not a carrier. This report has already split the broader category of all projected/displayed media into more useful groupings by removing the *microforms*. Surely *microfilm* and *microfiche* fit the definition of being projected/displayed just as surely as the images on film and slides must be magnified to be viewed. Retaining the separate category of *microform* is acknowledgement of how materials are used (it's the repugnance factor). We would argue it is finally time to split the remaining materials into categories that better match how the materials are actually used: *video* and *film*. *Film* is viewed by passing light through a translucent material. *Video* is viewed by optically or magnetically reading encoded data. These fundamental differences are critical to how the materials are used. Bluntly put: a motion picture on film and a motion picture on DVD are just as fundamentally different as a book in print and a book on microfilm. Why make the distinction for text but not for moving image?

In addition, there are technical errors that will need to be cleaned up, such as the way NTSC, PAL and SECAM are listed as carrier terms under *projected*. They are characteristics of specific carriers (i.e. DVD video, videocassette, videodisc) but not the carriers themselves. In addition to these, region codes should be added under DVD-video, for example: region 1, region 2, all region. *Video with audio description* does not denote a carrier. It describes two types of work, and that the resource was designed to be of particular use to deaf and hard-of-hearing people.

**Tactile:** *Tactile* describes a "script", not a carrier. Tactile resources can have paper, plastic, etc., carriers.

**Three-dimensional:** *Three-dimensional* describes the dimensionality of a work, expression, or manifestation, not a carrier. As the specific term list shows, three-dimensional resources can have any type of carrier. These terms do not denote types of carrier: *clothing, coins, diorama, doll, exhibit*. These terms describe type of carrier along with non-carrier information: *ceramic vase, flint arrowhead*.