

Of apples, glue, and Occam's razor

A critical look at the theory of aggregates

History of the Aggregates Model

1. IFLA Working Group on Aggregates: Final Report (2011)

it should be noted that there was disagreement among the group

➡ *alternative approach championed by part of the group in appendix B*

➡ *“winner model” was widely discussed and also criticized (e.g. RDA list)*

2. IFLA LRM (2017)

adopts the Working Group’s “winner model” mostly verbatim

➡ *only small changes, but no reappraisal of the matter*

3. Official RDA

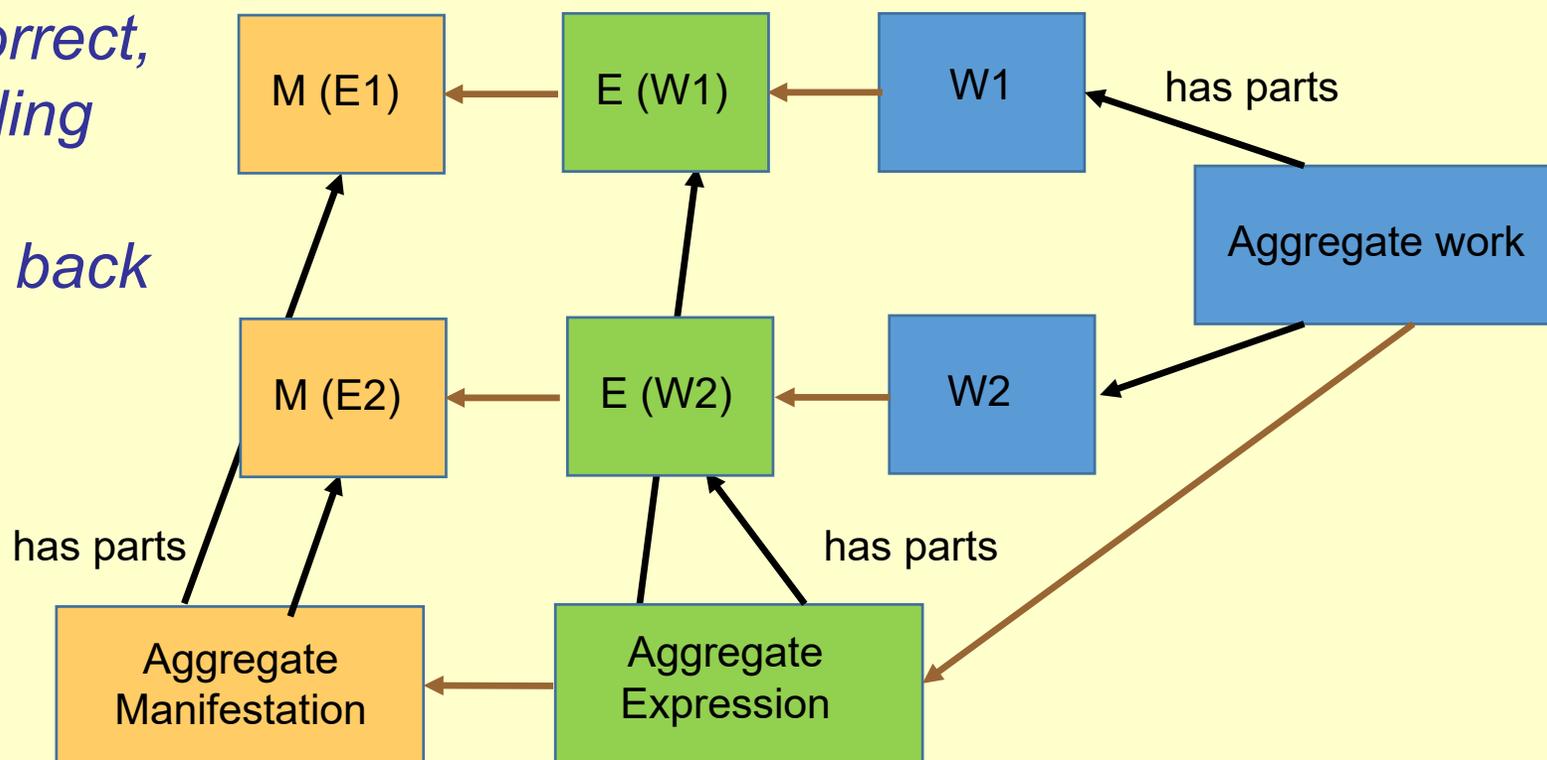
spells out the model for practical implementation

“Loser Model”

- **Aggregates exist on all WEMI levels**

“an aggregate entity is the “whole” in a “whole/part” relationship with two or more components (parts)” (WG report, p. 19)

➔ *seems intuitively correct, but there are modeling problems in certain situations (I’ll come back to this at the end)*



“Winner Model”

- **Aggregate manifestations only**

introduction of “aggregating works” and “aggregating expressions”

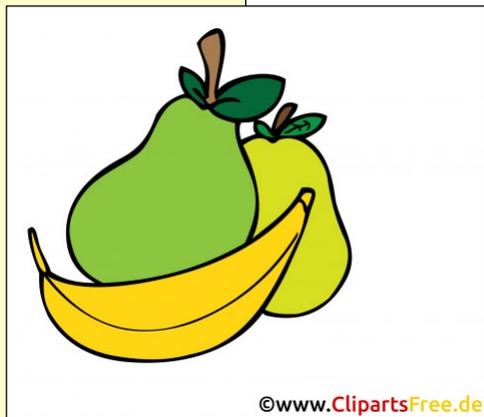
However, modeling an aggregate simply as an embodiment of discrete expressions may fail to recognize the creative effort of the aggregator or editor. The process of aggregating the expressions itself is an intellectual or artistic effort and therefore meets the criteria for a work. In the process of creating the aggregate manifestation, the aggregator produces an *aggregating work*. This type of work has also been referred to as the glue, binding, or the mortar that transforms a set of individual expressions into an aggregation. This effort may be relatively minor—two existing novels published together—or it may represent a major effort resulting in an aggregate that is significantly more than a sum of its parts (for example an anthology). An aggregating work is not a discrete section or even necessarily an identifiable part of the resulting manifestation and does not contain the aggregated works themselves.

Final report, p. 5

The intellectual effort of creating *aggregates*, such as selection and arrangement decisions, is a distinct intellectual or artistic creation and therefore is itself a *work*; a special type of *work* identified as an *aggregating work*. *Aggregating works* can be thought of as the frame, glue, binding, or the mortar that transforms the set of individual *expressions* into an *aggregate*.

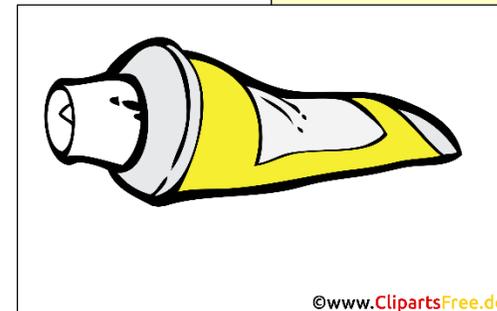
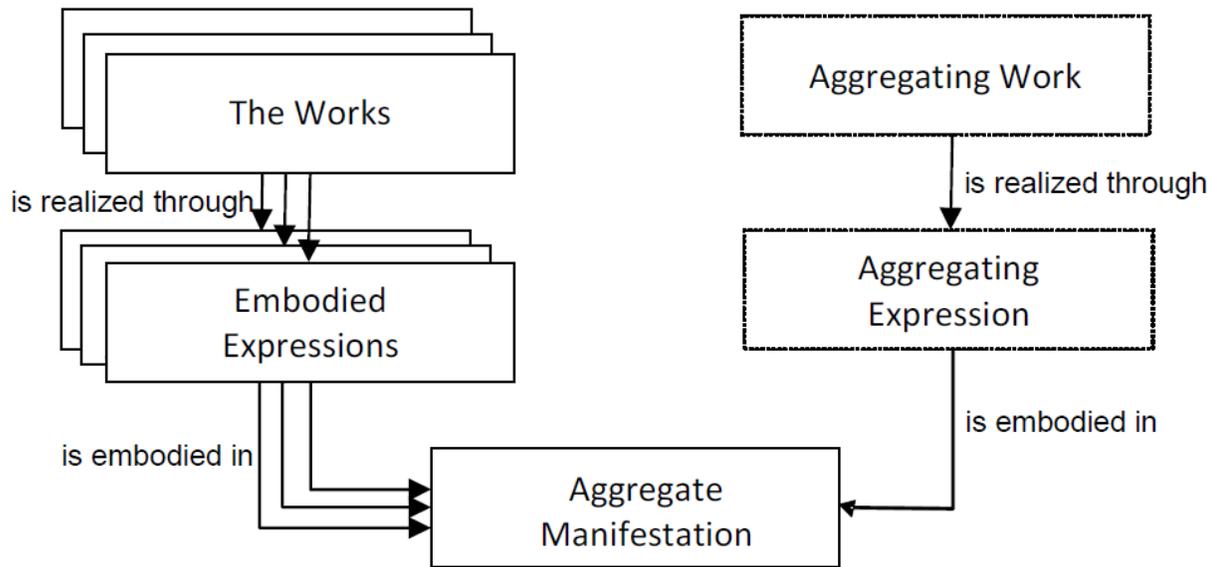
Final report, p. 6

The modeling of aggregates as a manifestation embodying multiple expressions is simple and straightforward; works and expressions are treated identically regardless of their form of publication or the physical manifestation in which they are embodied. An expression may be published alone or it may be embodied in a manifestation with other expressions. The general model for aggregates is shown in figure 3. A new entity is created for the aggregate manifestation which embodies n individual expressions. Any of these expressions may also be embodied individually in a non-aggregate manifestation.



©www.ClipartsFree.de

“Apple works”
(my metaphor for “normal” works)



©www.ClipartsFree.de

“Glue work”
(my metaphor for an “aggregating” work, inspired by the WG)

Figure 3. General Model for Aggregates.

Final report, p. 5

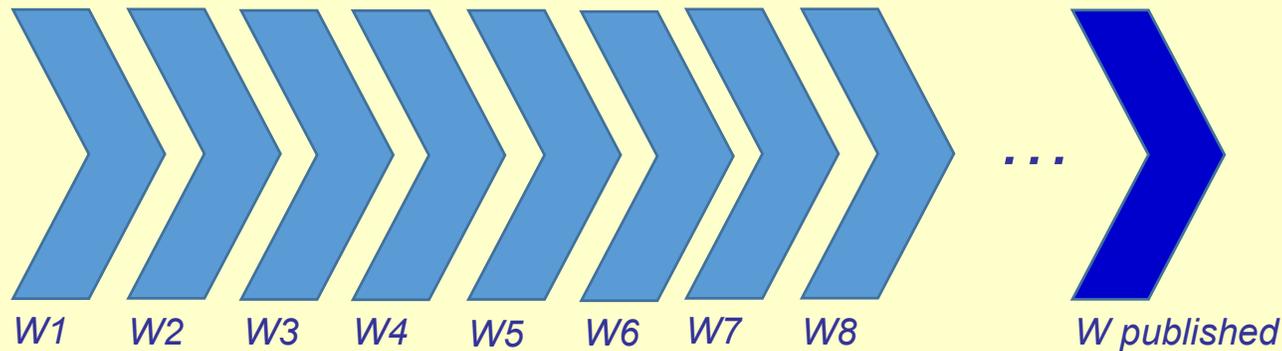
Are “glue works” really works?

- **Definitions of a work**
 - *Work in FRBR: “a distinct intellectual or artistic creation”*
 - *Work in LRM: “the intellectual or artistic content of a distinct creation”*
- **Slightly different definitions and explanations of “aggregating work”**
 - *WG₁: “glue, binding, or mortar” (metaphors)*
 - *WG₂: “effort ... such as selection and arrangement decisions” (p. 6)*
 - *LRM₁: “the concept or plan for the selection, assembly and ordering” (p. 21)*
 - *LRM₂: “the selection and arrangement criteria” (p. 94)*
 - *RDA: “a plan to select and arrange two or more expressions”*
 - ➔ *difficult to think of this as distinct and as a creation; rather, the creation is what we used to think of as the aggregate work (†), with the “glue” only being a means to this end [Note: † used here for “deceased”]*
 - ➔ *but if this is a distinct creation and therefore a work, why isn’t the plan to write, e.g., a textbook or a novel also a work in its own right?*

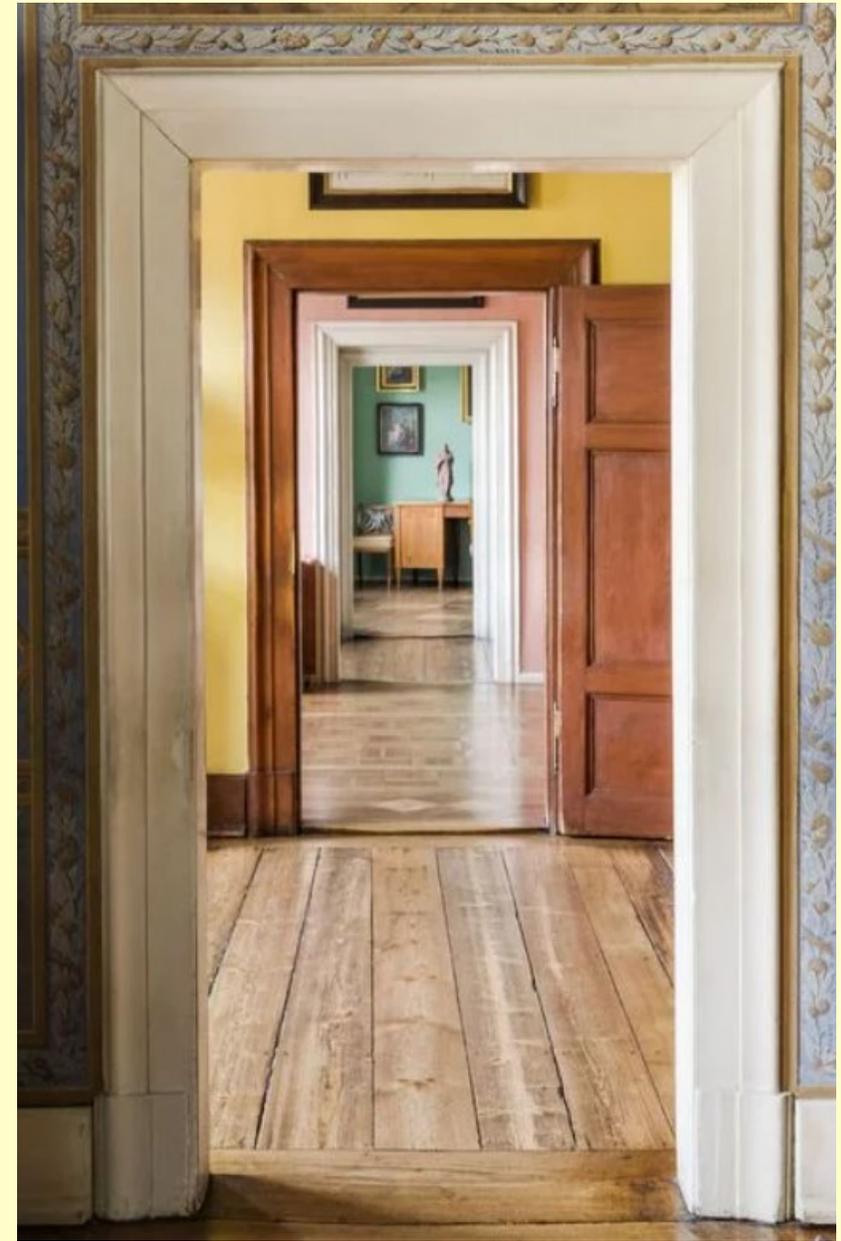
Odd abundance of “glue works”

- **Example: an edited collection on AI**
 - *starts with a vague idea of the editors*
 - *draft list of contributors, various changes*
 - *draft order of essays, readjustings*
 - *each text is edited in many ways*

➔ *each step represents a new “glue work”, so that you end up with a weird chain of works*



➔ *compare with “apple works”: similar steps are considered to be on expression level*



The Goethehaus at Weimar

Ok, let's treat the glue like a work

- **Elements of the “glue work”**

if we take the theory seriously, we need to distinguish it from the result (so we cannot use the title of the compilation “Algorithms of power”)

- *preferred title: devised, e.g. “Plan for edited collection on Artificial Intelligence”*

- *creators: the editors of the collection (but in many cases unknown)*

-  *AAP: Miller, Sue, 1980-. Plan for edited collection on Artificial Intelligence*

- *date of work: year of publication (as earlier works in the chain unpublished)*

- *category of work: aggregating work*

- *nature of content: not applicable*

- **And what about different creators aggregating the same expressions?**

e.g., two editions of “Alice in Wonderland” with the illustrations of John Tenniel

- *identical plan of selection and arrangement should be the same “glue work”*

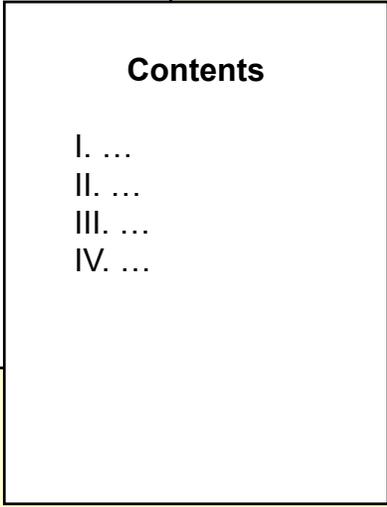
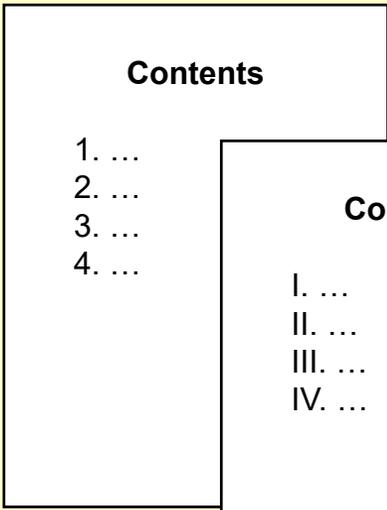
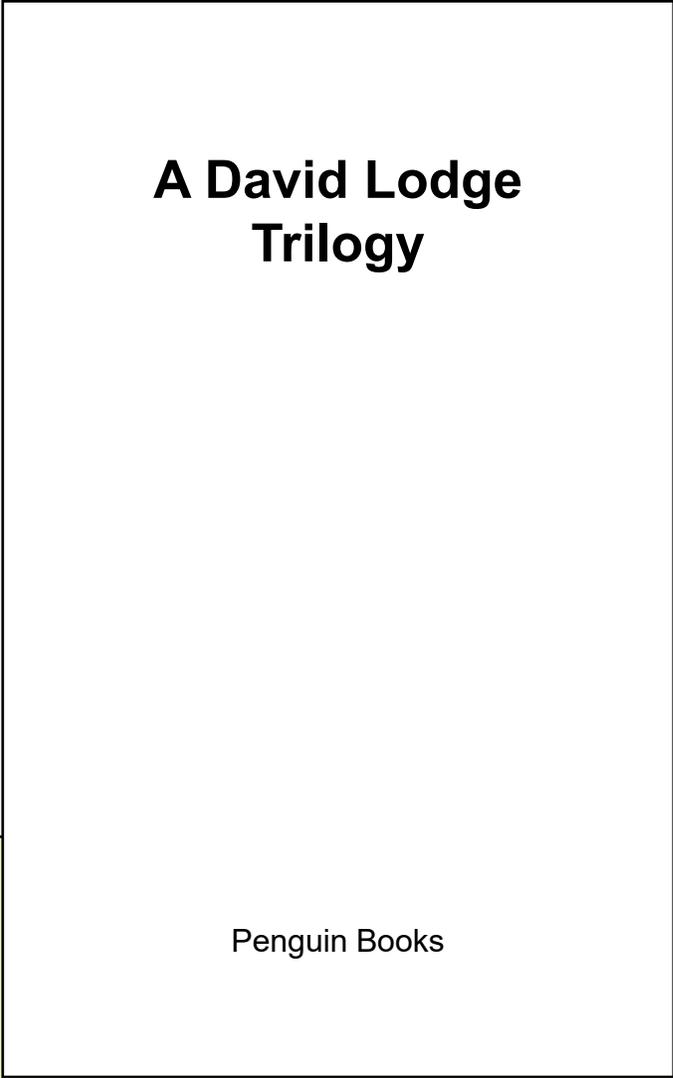
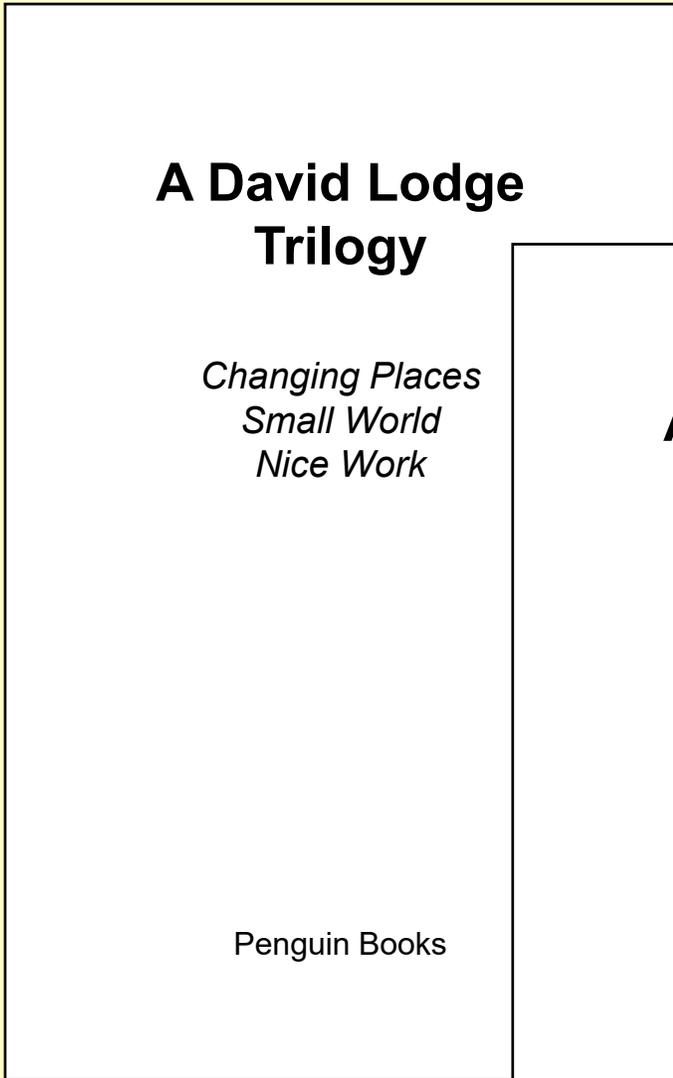
- *but different creators normally lead to different works*

The mystery of the “glue expression”

- **Definitions of expression**
 - *FRBR*: “intellectual or artistic realization of a work in the form of alpha-numeric, musical, or choreographic notation, sound, image, object, movement, etc.”
 - *LRM*: “a distinct combination of signs conveying intellectual or artistic content”
- **Definitions and explanations of “aggregating expression”**
 - *WG and LRM*: no definition or explanation given
 - *RDA*: gives both a definition and an explanation (88.69.69.51)

An *aggregating expression* realizes the plan of an aggregating work to select and arrange expressions that are embodied by an aggregate.

The content of an aggregating expression is what is embodied by an aggregate that excludes the content of the expressions that are aggregated. For example, a title page and page headings and numbering in a printed volume aggregate may be treated as the content of the aggregating expression.



➔ *seems to imply that differences like this should be treated as different aggregating expressions and, therefore, different aggregating works*

➔ *intuitively: these are differences on the manifestation level*

Ok, let's treat it like an expression

- **Elements of the “glue expression”**
 - *preferred title of the expression: ???*
(If we accept the RDA explanation, then maybe it would be the title proper of the manifestation? But then again, this identifies the result and not the realization of the plan!)
 - *creators: probably identical to the creators of the “glue work”*
 - *content type: probably not applicable*
 - *language and script: not applicable, unless we accept the RDA explanation (then we might use this according to the language and script of the title page)*
- ➔** *on the whole, a difficult and not very satisfactory description*

Self-assessment of the Working Group

The proposed approach was tested and evaluated with a wide variety of different manifestations that were either aggregates or possibly could be consider aggregates. The conclusion was that the proposed approach: (1) preserves the integrity of expressions and works, (2) is relatively easy to understand and apply, and (3) is consistent with the FRBR model.

Final report, p. 7

- **“relatively easy to understand”**

I find it absolutely mind-blowing, especially

- *the treatment of decisions, criteria or plans as works in the first place*
- *what an aggregating expression really *is* (can't be the resulting aggregate)*
- *the putting together of two very different kinds of things (“apples” and “glue”)*

- **“consistent with the FRBR model”**

oddly, variations which for “apple works” would be on expression level are considered to be on work level, which leads to the equally odd WE lock

➡ *e.g. a translation of an “apple work” is a new expression, but a translation of an edited collection is a new work*

Official Toolkit

Describing a manifestation that embodies two or more expressions

A manifestation that embodies two or more expressions of distinct works is an *aggregate*. An aggregate also embodies the one and only *aggregating expression* of an *aggregating work*.

CONDITION

A manifestation is an *aggregate*.

OPTION

Record information about one or more of the expressions that are embodied using Manifestation: [note on manifestation](#).

OPTION

Relate the manifestation to the aggregating expression using Manifestation: [expression manifested](#).

OPTION

Relate the manifestation separately to one or more of the expressions that are aggregated using Manifestation: [expression manifested](#).

OPTION

Relate the manifestation to the aggregating work using Manifestation: [work manifested](#).

OPTION

Relate the manifestation separately to one or more of the works that are realized by the expressions that are aggregated using Manifestation: [work manifested](#).

OPTION

Relate the manifestation separately to one or more of the creators of one or more of the expressions that are aggregated using Manifestation: [contributor agent to aggregate](#).

OPTION

Relate the manifestation to aspects of the content of one or more of the expressions that are aggregated using any of the following elements:

- Manifestation: [accessibility content](#)
- Manifestation: [colour content](#)
- Manifestation: [illustrative content](#)
- Manifestation: [sound content](#)
- Manifestation: [supplementary content](#)

35.26.22.03

- **Following these links there is no special guidance**
e.g. “Manifestation: expression manifested” tells you to record this element as title of expression, access point, identifier or IRI, but there are no special options for, e.g., the preferred title or the AAP of an aggregating expression
- **Describing an expression of an aggregating work** (80.89.44.19)
 - *options for giving information about the creator(s) (creator agent of expression)*
 - *options for giving information about the aggregated expressions*
- **Describing an aggregating work** (86.91.46.53)
 - *options for giving information about the creator(s) (aggregator agent)*
 - ➔ *description is in terms of relationships only, whereas no guidance is given about attributes of aggregating works and expressions, especially the question of the preferred title and the AAP*
 - ➔ *this gives a lot of “wiggle room” (which presumably is by intention)*

New LC-PCC practice (not yet implemented)

Official RDA Toolkit

LC-PCC Metadata Guidance Document

*LC-PCC Metadata Guidance Document "Aggregates"
(2022-09-26, p. 4)*

By the time of the publication of the new RDA, the dust had settled on the competition between the two models and the aggregates model had emerged victorious. Thus, this guidance represents a completely different way of thinking about how content is embodied in a manifestation.

While this is a new way of thinking, the results should look more or less identical to the data produced using original RDA, save for the use of new relationships. These are listed below.

Original RDA	Official RDA
RD: container of (work)	Work: part work . Do not use for aggregates. Instead use Manifestation: expression manifested .
RD: container of (expression)	Expression: part expression . Do not use for aggregates. Instead use Manifestation: expression manifested .

- **LC-PCC Task Group on Aggregates**
2020 report with a very thorough discussion of problems and options

- **One actual change in cataloging practice**
use of different relationship in 700 for aggregated works/expressions (in order to avoid whole/part terminology)

New LC-PCC practice (not yet implemented)

100+240 = expression manifested

Example 1

A collection of short stories

The [Manifestation: expression manifested](#) for the aggregating expression is recorded using an access point for the expression which is split between the 100 and 240 fields. The 100 field is also identifying the creator of the primary content.

```
100 1# $a Healy, Dermot, $d 1947-2014, $e author.
240 10 $a Short stories
```

```
245 14 $a The collected short stories / $c Dermot
Healy ; edited, with an introduction, by Keith
Hopper & Neil Murphy.
```

p. 15

```
245 00 $a Algorithms of power / $c edited by Sue Miller.
```

- **Change in wording only**
redefining what used to be the AAP for the aggregate work (†) as the AAP for the aggregating expression

Cf. this prophetic statement by Preston Salisbury on RDA-L (2019-05-30)

“However, for practical cataloging purposes, one is still allowed to behave as if there are aggregate works. (...) One is allowed to treat the “aggregating work” in the same way as “aggregate works” have been treated, as long as “part-of” language is avoided. (...) the difference between “aggregating” and “aggregate” is so small as to be meaningless.”

- **Pro's**

- *catalogers only need to learn little changes for their practical cataloging*
- *no big divergence between old and new data*
- *users aren't presented with irrelevant or weird-looking information (like my suggested AAP with "Plan for ... ")*

- **Con's**

- *if the model has been understood and is taken seriously, this solution is very difficult to reconcile with it, as it seems to mix up the new aggregating work and the former aggregate work (†), which also makes it difficult to teach*
- *it may be unclear to catalogers why they need to learn about a difficult theory if there are almost no practical changes*
- *it could be interpreted as "cheating"*
- *it's unclear whether the WE lock is adhered to (if it isn't, this would be a violation of the model)*

The WE lock and Occam's razor

- **WE lock for collection aggregates**

e.g. a second, revised edition belongs to a different aggregating work

245 00 \$a Algorithms of power / \$c edited by Sue Miller.
250 ## \$a Second, revised edition.

➔ *this solution violates the model, as it implies the same aggregating expression (and, therefore, the same aggregating work) as the first edition*

- **Possible solutions**

- *differentiate in 130 and maybe add the level of “work group” for collocation*
- *don't change the practice, but only the wording (“work group” instead of “work”)*

➔ *both lead to new entities which don't really seem necessary (and in my opinion undermine the original FRBR concept of a work)*

Occam's razor:

“Entia non sunt multiplicanda praeter necessitatem.”
 (“Entities are not to be multiplied without necessity.”)

My Conclusion

- **Inherent problems with the present aggregates model**
not restricted to MARC implementation, e.g., the rift between “apples” and “glue” and the need for an additional layer (work group) don’t go away with BIBFRAME
 - **Maybe it’s time to take another look at the “loser model”**
 - *fits much better with the intuitive understanding*
 - *complies with the principle of representation (e.g., a 2nd ed. of an edited collection is not presented by editors and publishers as a new work)*
 - *main criticism: doesn’t capture when a single work is later published as part of a collection, implies that the work is *always* part of this aggregate work (†)*
 - *similar problem with the “winner model”: doesn’t capture that sometimes works only get created to be part of an aggregate work (†), as in an edited collection*
- ➔ *personally, I believe the “loser model” could be adapted adequately and would then be a better basis for the description of aggregates*

References

IFLA Working Group on Aggregates (2011): Final report, <https://cdn.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/frbrrg/AggregatesFinalReport.pdf>

LC-PCC Metadata Guidance Document: Aggregates (2022-09-26), <https://www.loc.gov/aba/rda/mgd/mg-aggregates.pdf>

LC-PCC Task Group on Aggregates in Beta RDA Toolkit (2011): Final report, part 1 (2020-08): <https://www.loc.gov/aba/pcc/taskgroup/Aggregates-TG-Final-Report-Part-1.pdf>, part 2 (2021-02): <https://www.loc.gov/aba/pcc/taskgroup/RDA-Aggregates-TG-Final-Report-Part%202.pdf>