

# Non-RDA Entities Task Group Interim Report

August 31, 2020

The Task Group (TG) on coding non-RDA entities has met five times since July 30, 2020.

The TG chose to focus first on the entity type vocabulary, and to defer consideration of coding for descriptive conventions until a later date. Because the German National Library (DNB) offered a clear precedent for coding entity types in a library authority file, the TG spent some time studying the German vocabulary and the MARC proposal submitted by DNB for the authority 075 field. In addition, the TG contacted DNB for background information on the creation of their vocabulary.<sup>1</sup> The TG also studied examples from the Getty Art & Architecture Thesaurus.

The DNB experience provided some useful lessons for the TG. For example, our DNB colleagues emphasized the need for an ongoing editorial process, and for clear definitions. However, the TG determined that the DNB vocabulary would not itself be suitable for PCC use. This was due in large part to the fact that its hierarchies do not map cleanly to RDA: for example, fictitious characters are a subclass under Person. As we learned from our DNB colleagues, many of the characteristics of their vocabulary reflect historical practice in their communities.

The focus of the group then turned to considering alternative models for a vocabulary to identify the entity types represented in the Name Authority File. TG members developed several models for consideration.

- Model A proposed adopting the German vocabulary, or relevant portions of it, with translations as needed. While this option was the simplest from a practical viewpoint, the TG has rejected it for reasons given above.
- Models B and D used the RDA vocabulary for entities covered by RDA and created a parallel vocabulary for entities RDA does not cover. Model D differed from Model B in proposing an upper level “proto-Agent” entity that subsumes both RDA and non-RDA entities.
- Model C created a separate vocabulary covering desired entity types and mapped subsets of the vocabulary to RDA entities.
- Model E sought to categorize the distinctions among real and fictitious entities, and entities not falling into either category.

---

<sup>1</sup> The TG wishes to thank the DNB’s Esther Scheven and Reinhold Heuvelmann for providing background to the DNB project.

- Model F presented a high-level vocabulary, preferring multiple type assertions (e.g. treating fictitiousness as a separate assertion) over embedding types in a deep hierarchy.

In the course of discussing these models the TG came to several provisional conclusions about the approach it should take.

- The vocabulary should have a narrow scope focused on cataloging and authority use cases. Where more granular entity type assertions (such as breed of dog or type of ship) are needed, these could be accomplished using other vocabularies, or asserted outside the NAF altogether, for example in Wikidata. The TG will review existing cataloging documentation (e.g. the current RDA and DCM Z1) to identify relevant entity types.
- Rather than try to anticipate all possible needs, the vocabulary should address known use cases and allow for the possibility of further development at a future date.
- As foreseen by the charge, issues surrounding pseudonyms will require careful consideration, and will largely fall outside the scope of the present TG. The question of attribution raises larger modeling issues that go beyond entity type.
- While the proposed model may raise questions about some aspects of existing practices, the TG affirmed that changes to authority practice beyond populating the 075 and 040 fields are beyond its scope.

In the course of its meetings, the task group has identified several questions for PoCo.

1. Can PoCo confirm that the scope of the task group is limited to types of entities that are currently established in the Name Authority File?
2. Would PoCo consider supporting a request to LC to host an entity vocabulary on the id.loc.gov site?
3. Would PoCo consider establishing an ongoing editorial process to maintain the vocabulary?