Towards a Vocabulary for Mapping Quality Assessment

Alex Randles [0000-0001-6231-3801], Ademar Crotti Junior [0000-0003-1025-9262] and Declan O'Sullivan [0000-0003-1090-3548]

ADAPT Centre, Trinity College Dublin, Dublin 2, Ireland
{alex.randles,ademar.crotti,declan.osullivan}@adaptcentre.ie

Abstract. This paper presents a vocabulary for expressing information related to the assessment, refinement and validation of mappings called the Mapping Quality Vocabulary.

1 Introduction

Oftentimes, RDF datasets are generated by converting non-RDF resources to RDF in a process called 'uplift'. The uplift process commonly involves the definition of mappings, which allows one to declaratively express the transformations needed to convert non-RDF source data to RDF [1]. Another use for mappings is found when relating and interlinking those RDF datasets. The creation of such mapping definitions is a complex time-consuming task, involving various quality related activities in which mappings are iteratively refined until they satisfy its stakeholders expressed requirements. While approaches for assessing the quality of mappings have been proposed [2, 3], a vocabulary for describing such processes is still lacking. To tackle this problem, we present the Mapping Quality Vocabulary (MQV) which aims at enabling quality metadata and provenance information relating to the assessment and refinement of mappings to be captured and published.

2 Mapping Quality Vocabulary

The proposed Mapping Quality Vocabulary[1] provides a vocabulary for expressing information relating to the quality assessment, refinement and validation of mappings. The goal is to make this information easier to publish, exchange and consume. Our proposed model is separated into three stages: **assessment**, **refinement** and **validation**. **Fig. 1** provides a general overview of the core components of the MQV model.

Mapping quality assessment. In this stage, one or more mapping documents (`mqv:MappingDocument`) are assessed. An assessment activity is captured through the `mqv:MappingAssessment` class. A mapping assessment activity may have

[1] Mapping Quality Vocabulary Specification available at `https://alex-randles.github.io/MQV/`

quality requirements, which are captured through the `mqv:QualityRequirement` class. We declaratively capture such information as we foresee and allow for quality validation activities to be executed in later stages of the process. The model uses this information to generate a mapping validation report (`mqv:MappingValidationReport`). Each violation identified is then represented with the `mqv:MappingViolation` class.

Mapping quality refinement. This stage involves capturing mapping refinements which are executed on the mapping to remove quality violations. Each metric described using our model may have multiple refinements (`mqv:MappingRefinement`) depending on the quality aspect being measured. The refinement executed in the mapping is associated with the identified violation through the property `mqv:wasRefinedBy`.

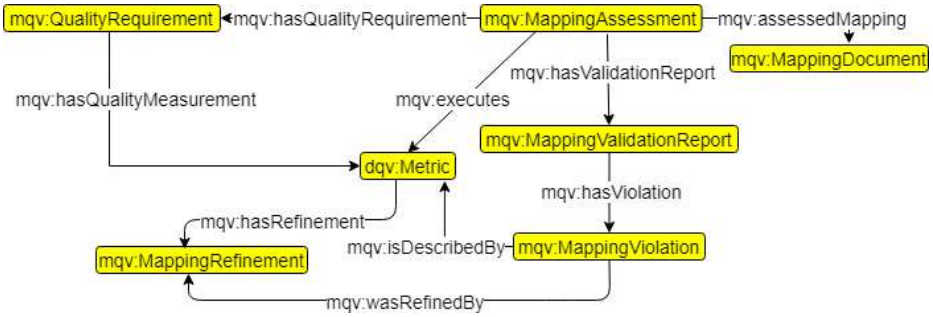

Fig. 1. Core components of the MQV model.

Mapping quality validation. Finally, our model provides quality information on the original mapping being assessed, and the mapping which has been refined in the process. As mentioned, each mapping assessment process may have quality requirements which can be validated at this stage.

3 Conclusion

The MQV model allows mapping quality information to be captured in a machine-readable format which allows it to be easily interpreted and processed by agents. Publishing this information as metadata is expected to improve the trustworthiness of the dataset, as well as encouraging the reuse and maintenance of those mappings.

References

1. Crotti, A. et al.: An evaluation of uplift mapping languages. Int. J. Web Inf. Syst. 13, 4, 405–424 (2017). https://doi.org/10.1108/IJWIS-04-2017-0036.
2. Junior, A.C. et al.: Assessing the Quality of R2RML Mappings. In: SEM4TRA-AMAR@SEMANTICS. (2019).
3. Moreau, B., Serrano-Alvarado, P.: Assessing the Quality of RDF Mappings with EvaMap. In: 17th Extended Semantic Web Conference. (2020).