

Implementing and managing mappings for data transformation using SHACL Rules

Lucas Verhelst ¹

¹ *BIM-Connected, Netherlands*

Abstract

Large asset managing organizations that manage assets typically have multiple systems that store and use similar or related information. Data management is a growing challenge for these organisations, which is why some of them have started to develop common data models in the form of ontologies. Implementing and managing mappings for data transformation using SHACL rules, RML, and Liquid templates can be a powerful solution for creating a single source of truth from different source systems. This allows organizations to move away from software with hard-coded mappings and towards a data-driven platform. The software can then be generic, and the mappings to the common data model can be reused for multiple purposes. However, there are also some challenges to consider when implementing such a platform. One of these challenges is that there is currently little off-the-shelf software available to process the SHACL rules, which means that organizations are depending on a few implementations which have their issues.

Another challenge is how to give non-technical stakeholders insight into the more complex mappings. While SHACL rules, RML, and Liquid templates are powerful tools, they are difficult for non-technical stakeholders to understand. To address this challenge, organizations need visualization tools to present the mappings in a more user-friendly way. This is an essential challenge to overcome, as the stakeholders cannot trust the results of data integration if they can't see the details of this process. Not being able to see the mapping rules also makes them difficult to govern for subject matter experts.

Despite these challenges, there are several benefits to developing such a platform for an organization. As mentioned before, it allows organizations to move from a siloed landscape towards a single source of truth, where all data is stored in a standardized format and can be accessed by all stakeholders. By storing the mappings as data, transparency about the way data is transformed across the platform is improved. The mappings also provide an overview of the differences in the datamodels of the connected systems.

Implementing and managing mappings for data transformation using SHACL rules, RML, and Liquid templates can be a powerful way to achieve map different systems to a common data model. This method involves using Liquid or RML to translate the data from the source format to linked data in a schema that is close to the source schema and then using SHACL rules to map the data to the common data model. In this 2-step approach, the technologies complement each other by each providing functionality the other lacks.

In conclusion, implementing and managing mappings for data transformation using SHACL rules, RML, and Liquid templates can be a powerful solution for integrating data from multiple systems. While there are some challenges to consider, such as the lack of off-the-shelf SHACL rules processors and giving non-technical stakeholders insight into complex mappings, the benefits of such a platform, including moving towards a single source of truth and using open standards, make it a valuable tool for organizations.

Keywords

Data Integration, Knowledge Graphs, Ontologies, Data Transformation, Linked Data, Common Data Model, Governance, Open Standards

Proceedings LDAC2023 – 11th Linked Data in Architecture and Construction, June 15–16, 2023, Matera, Italy
EMAIL: lucas.verhelst@bim-connected.com (L. Verhelst)



© 2023 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)