OntoQAV: A Pipeline for Visualising Ontology Quality

Silvio Mc Gurk[1], Jeremy Debattista[2][*], and Charlie Abela[1]

[1] University of Malta,
`silvio.mcgurk.15@um.edu.mt`, `charlie.abela@um.edu.mt`
[2] ADAPT Centre, School of Computer Science and Statistics,
Trinity College Dublin, Ireland
`debattij@scss.tcd.ie`

Abstract. Ontological Quality has been the subject of extensive research due to the importance of ensuring that a system's underlying ontologies are fit for purpose. Understanding quality problems may not be straightforward, depending on the ontologies' magnitude and complexity, the nature and extent of the problems, and the metrics used in its quality assessment. In this poster paper we present an innovative pipeline[3] linking together a quality assessment framework (Luzzu) and an ontology visualisation framework (WebVOWL) in order to establish an ecosystem whereby knowledge engineers can assess and interactively understand quality problems within concepts and properties in ontologies.

Keywords: Ontology Quality, Quality Framework, Quality Assessment, Quality Visualisation, Ontology Visualisation

1 Introduction and Context

The use of ontologies has become widespread across many domains (e.g. biological, geographical, government, etc...) as they provide the means for sharing concepts and data among different organisations [5]. Together with the RDF standard, the idea behind ontologies was to solve data interoperability problems. However, choosing a fit-for-use ontology for a system might not be the simplest task. Research carried out by [4], amongst others, resulted in a number of metrics being proposed to help identify quality problems. Nonetheless, without the right tools, ontology stakeholders still encounter difficulties when choosing the right ontology for the task at hand.
Inspired by the progress achieved on Linked Data quality frameworks [1,4] and ontology visualisation frameworks [2], we aim to exploit these state-of-the-art frameworks in order to address a missing niche in ontology quality, that is,

---

[3] Demonstration: http://github.com/silviomcgurk/OntoQAV

assisting system engineers in finding the fit-for-use ontology. In this poster paper we present OntoQAV, a pipeline integrating Luzzu [1], a generic Linked Data quality assessment framework, and WebVOWL [2], a widely used tool to represent RDF-based ontologies in a visual format. The rational behind this pipeline is to create an ecosystem with the aim of allowing stakeholders to assess ontologies on various quality metrics and to visualise any identified problems. We aim to provide the possibility of assessing multiple ontologies at once and presenting a comparative visualisation and summary of the quality problems within the ontologies being assessed. The main contributions of this pipeline are: (1) implementing ontology quality metrics for Luzzu identified in our previous work [3]; and (2) implementing a plug-in for WebVOWL which takes as input a quality problem report from Luzzu and displays the problems in WebVOWL. Metrics for Luzzu and the WebVOWL plug-in are available in a public repository, along with instructions and links for a demonstration of the pipeline (available at http://github.com/silviomcgurk/OntoQAV).

2 The OntoQAV Pipeline

Figure 1 depicts OntoQAV. The pipeline follows a three-step workflow. In the first step (**Step 1**), a user selects the ontology that needs to be assessed and the relevant quality metrics that represent *fitness for use.* Following that, (**Step 2**) the Luzzu quality assessment framework is initialised and the assessment gets underway [1]. The assessment provides the ontology engineer with two results: (1) the quality metadata, which is preserved on the Web of Data for further quality-based tasks (e.g. filtering ontologies based on different quality criteria); and (2) the quality problem report, which is preserved on the Web as Linked Data for future use (possibly by other interested parties). The quality problem report is then converted to JSON-LD format and fed to WebVOWL, the third step (**Step 3**).

Therefore, the result is an augmented visualisation which gives a graphical representation of the nodes and links within the ontology, with an additional visualisation layer of the problems identified by the quality assessment framework.

3 Visualising Ontology Quality Problems

Luzzu [1] and WebVOWL [2] are the two tools chosen in our proposed pipeline to demonstrate a proof-of-concept for visualising ontology quality. Furthermore, in the future, we aim to make this pipeline generic, implementing and making use of mechanisms and vocabularies that facilitates the data exchange between existing and future tools. Luzzu is an extensible Linked Data quality assessment framework. It also provides quality metadata and problem reports that can be leveraged in semantic-driven frameworks for other tasks. WebVOWL is an ontology visualisation tool, representing concepts and properties of ontologies in a
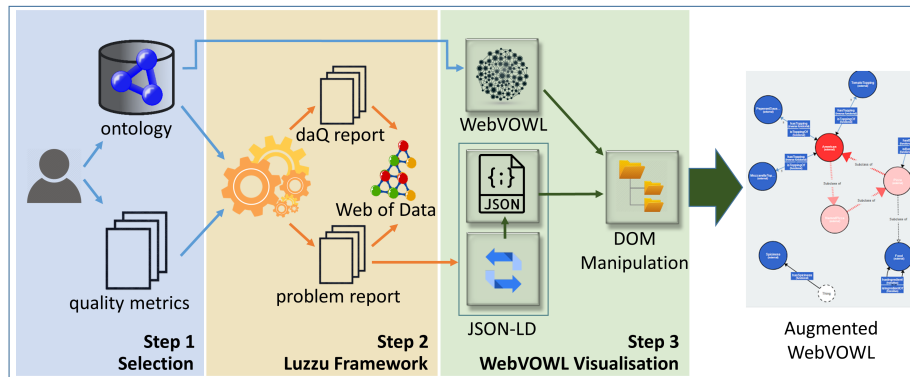
Fig. 1: Proposed OntoQAV Pipeline

way that can be easily understood. Upon completion of the quality assessment for a given ontology, Luzzu provides a Linked Data structured problem report (cf. Listing 1.1 for a snippet) highlighting problematic concepts and properties for each assessed metric. This problem report is then converted into a JSON-LD serialisation and used during the modification of the WebVOWL DOM objects to represent the problematic concepts.

```
ex:QualityProblem a qpro:QualityProblem ;
    qpro:isDescribedBy ex:CyclesInOntologyMetric ;
    qpro:problematicThing [
      a rdf:Statement ;
      rdf:object ex:Pizza ;
      rdf:predicate ex:subClassOf ;
      rdf:subject ex:NamedPizza ] .
```

Listing 1.1: Snippet from Luzzu Problem Report (Turtle)

Together with the JSON-LD serialisation of the problem report, the assessed ontology is loaded into WebVOWL for its visualisation. Following the loading of the ontology, the user is given a choice to view the problematic concepts, upon which the proposed rendering of quality visualisation is triggered. At this stage, the pipeline plug-in interacts with the rendered visualisation of the ontology and modifies it through the browser's Document Object Model (DOM) to augment the visualisation with quality information. Shading of the red colour has been selected to represent quality issues within an ontology. For every problem identified by Luzzu, the plug-in gives a red shade to the problematic components (nodes, properties or relationships). Components that fail more than once (with different metrics) will have a darker shade. As a result, the shading of red from light to dark colour indicates the extent of possible quality issues of the respective component. Additional information regarding the quality problems of concepts and properties of the visualised ontology are shown to the user in the WebVOWL sidebar. Figure 2a shows the visualisation of the pizza ontology rendered in We-

bVOWL. Once the quality visualisation plug-in is invoked, problematic concepts are highlighted in red (cf. Figure 2b). In this example, the pizza ontology has been assessed by the *Cycles in Ontology Metric* [3]. The assessment identified the concepts `American`, `NamedPizza`, and `Pizza` and properties `rdfs:subClassOf` violating this metric, and thus are highlighted in red.
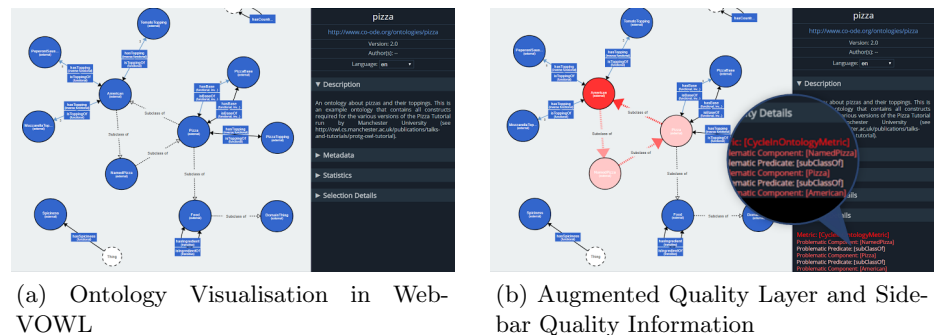

(a) Ontology Visualisation in Web-VOWL

(b) Augmented Quality Layer and Sidebar Quality Information

Fig. 2: Visualisation

4 Final Remarks and Future Work

In this poster paper we proposed an innovative pipeline that links together a quality assessment framework and visualisation tool in effort to establish an ecosystem to enhance the evaluation of ontologies from a quality perspective, providing an intuitive way of looking at various quality problems an ontology might have. Our future work includes a plan to do a comprehensive quality assessment on LOV[4] ontologies and make available their respective quality assessment, quality problem reports and augmented visualisation.

References

1. Debattista, J., Auer, S., Lange, C.: Luzzu - A Methodology and Framework for Linked Data Quality Assessment. Journal of Data and Information Quality. (2016).
2. Lohmann, S., Negru, S., Haag, F., Ertl, T.: Visualizing ontologies with VOWL. Semantic Web. 7, 399-419 (2016).
3. Mc Gurk S., Abela C., Debattista J.: Towards Ontology Quality Assessment. Joint proceedings of the MEPDaW 2017 and LDQ 2017. (2017).
4. Poveda-Villaln, M., Gomez-Perez, A., Suarez-Figueroa, M.: OOPS! (OntOlogy Pitfall Scanner!):. Intl. Journal on Semantic Web and Information Systems. (2014).
5. Ristoski, P., Paulheim, H.: Semantic Web in data mining and knowledge discovery: A comprehensive survey. Web Semantics: Science, Services and Agents on the World Wide Web. 36, 1-22 (2016).

---

[4] http://lov.okfn.org/dataset/lov/