R2[RML]-ChatGPT Framework

Alex Randles and Declan O’Sullivan

ADAPT Centre for Digital Content, Trinity College Dublin, Ireland
Motivation

➢ Creating high-quality mappings requires a lot of background knowledge

➢ Mapping quality issues can exponentially multiply into resulting data

Figure 1. Process of RDF creation
Motivation

- Existing quality assessment approaches often restricted to knowledge in mappings and used ontologies
- Diverse knowledge available in LLMs
- Support mapping engineers during mapping creation
  - What alternative concepts/ontologies are available?
  - Create sample instances which use this mapping concept?
  - Create a constraint to validate the range of this concept?
Design and Implementation of Framework
R2[RML]-ChatGPT Framework
Design of Framework

- R2[RML] mapping uploaded to framework

- Pre-processing involves retrieving distinct concepts and inputting into prompt templates into ChatGPT 3.5 turbo

- Post-processing extracts, validates and improves returned code
Workflow of Framework

**Figure 1:** Workflow of the R2[RML]-ChatGPT Framework
Workflow of Framework

**Figure 1:** Workflow of the R2[RML]-ChatGPT Framework
Figure 1: Workflow of the R2[RML]-ChatGPT Framework
Implementation of Framework

- Several Python libraries used to implement framework

<table>
<thead>
<tr>
<th>Library</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flask</td>
<td>Create web application for interface</td>
</tr>
<tr>
<td>SPARQLWrapper</td>
<td>Execute SPARQL queries</td>
</tr>
<tr>
<td>OpenAI</td>
<td>Communicate with ChatGPT</td>
</tr>
<tr>
<td>RDFLib</td>
<td>Syntax parser</td>
</tr>
</tbody>
</table>
Can you provide me with key information related to the ‘prov:generatedAtTime’ used in RDF/OWL Technology?

Can you provide the ‘rdf:type’ value for the ‘prov:generatedAtTime’ concept defined in an ontology used in RDF/OWL Technology?

**Figure 4:** Screenshot of Implementation displaying information related to concepts in a mapping
Implementation of Framework

Response from ChatGPT

Code exportation and validation

Figure 5: Screenshot of code validation and exportation available on the framework
Experimentation
Validating Syntax and Semantics
Experiment Overview

➢ **RQ1:** To what extent will ChatGPT produce semantically correct data for certain values in a declarative uplift mapping (e.g. type, domain, range and label)?

➢ **RQ2:** To what extent will ChatGPT produce syntactically correct RDF data and SPARQL queries?

**Figure 6:** Overview of Activities involved in the Experimentation
Testing Semantic Correctness (RQ1)

- 4 ontology terms (type, domain, range and label) retrieved for concepts in mappings

- Returned terms inserted in SPARQL ASK queries

- Queries executed on namespace ontology for comparison
Semantic Correctness (RQ1) Results

➢ Domain and Range scored similar

➢ Type scored slightly worse with the name of tested concept returned in some cases

➢ Label scored worst with inferences resulting in incorrect results

Figure 7: Results of correctness of each ontology term tested
Testing Syntactic Correctness (RQ2)

- 150 files containing Instances, SPARQL and SHACL generated

- Syntax validated using RDFLib parsers

- Invalid syntax improved by framework
  - Regular expressions applied on parser output
  - Added missing prefixes
  - Syntax validated again
Syntactic Correctness (RQ2) Results

- Mean score of 42 (84%) out of 50 files correct
- SHACL constraints scored best which could be due to less prefixes
- Post-processing resolved most (14 out of 18) syntax issues

*Figure 8:* Results of occurrences of syntactic correctness for each category tested before (left) and after (right) post-processing by the framework
Future Work

➢ Usability testing of framework

➢ Testing of other ontologies

➢ Extending support to interlink mappings

➢ Comparison of results with other LLMs and ChatGPT versions
Conclusion

➢ Labels scored worst with some inferences (e.g. rdfs:rest -> rest of list)

➢ Most common syntax problem was missing prefixes (14 out of 18)

➢ Level of ontology documentation could impact scores (rdfs better than skos)

alex.randles@adaptcentre.ie

CALL TO ACTION
Get involved in evaluating the VRTI-KG Explorer in coming months ...
... no technical or historical expertise needed!