W3C Community Group
Knowledge Graph Construction

David Chaves-Fraga
http://w3id.org/kg-construct
W3C Community Group - Knowledge Graph Construction

176 participants (~25-30 active)

Bi-weekly meetings

http://w3id.org/kg-construct

http://github.com/kg-construct
Towards the RML standardization

● Five specs:
  ○ RML-Core: Schema transformations
  ○ RML-IO: Source and target
  ○ RML-CC: Collection and containers
  ○ RML-FNML: Data transformation functions
  ○ RML-star: RDF-star
  ○ **RML-views: Complex data pipelines (ongoing)**

● Modular approach

● Unification of prefixes → [w3id.org/rml/](https://w3id.org/rml/)

● The RML portal → [https://w3id.org/rml/portal](https://w3id.org/rml/portal)
RML-Core: Schema transformations

- Maintains R2RML basic structure
- **Dynamic** generation of:
  - Language tags
  - Data types
- Increased **flexibility** for join conditions
RML-IO: Data source and target

- **Extended input** data source description
- **Output data** description
- Leverage of existing *vocabularies* (SCAT, SPARQL-SD, VoID)
RML-FNML: Data Transformations

- Refines RML+FnO approach
- Reference connector between RML and the Function Ontology (FnO)
RML-CC: Collections and Containers

- Introduces generation of collections and containers
- Specifies how gather terms into a CC and manage them: to assign them a IRI or BN, manage empty CC, how the gathering is performed...
RML-star: RDF-star

- **Recursiveness** in mapping rules to generate quoted triples
- Applicable in **subject** and **object** position
- Asserted and non-asserted quoted triples
RML overview (2020-2024)
RML-Logical Views

Draft Community Group

**Latest published version:**
none

**Latest editor’s draft:**
https://w3id.org/kg-construct/

**Editors:**
Thomas Delva (Ghent Univ)
Anastasia Dimou (Ghent Univ)

**This Version**
https://kg-construct.github.io/r

**Previous Version**
https://kg-construct.github.io/n

**Website**
https://rml.io/

---

**Example record sequence**

<table>
<thead>
<tr>
<th>#</th>
<th>&lt;it&gt;</th>
<th>name.#</th>
<th>name</th>
<th>item.#</th>
<th>item</th>
<th>item.type.#</th>
<th>item.type</th>
<th>item.weight#</th>
<th>item.weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>{...}</td>
<td>0</td>
<td>alice</td>
<td>0</td>
<td>{&quot;type&quot;: &quot;sword&quot;, &quot;weight&quot;: 1500}</td>
<td>0</td>
<td>sword</td>
<td>0</td>
<td>1500</td>
</tr>
<tr>
<td>0</td>
<td>{...}</td>
<td>0</td>
<td>alice</td>
<td>1</td>
<td>{&quot;type&quot;: &quot;shield&quot;, &quot;weight&quot;: 2500}</td>
<td>1</td>
<td>shield</td>
<td>0</td>
<td>2500</td>
</tr>
<tr>
<td>1</td>
<td>{...}</td>
<td>0</td>
<td>bob</td>
<td>0</td>
<td>{&quot;type&quot;: &quot;flower&quot;, &quot;weight&quot;: 15}</td>
<td>0</td>
<td>flower</td>
<td>0</td>
<td>15</td>
</tr>
</tbody>
</table>
Complete set of resources per module

- Specifications
- OWL ontologies
- SHACL shapes for mapping validation
- Test cases
- Backwards compatibility

RML Ontology Modules

Here you can find the list of modules of the mapping language RML.

<table>
<thead>
<tr>
<th>Ontology</th>
<th>Serialization</th>
<th>License</th>
<th>Language</th>
<th>Links</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RML-Core</td>
<td>OWL</td>
<td>CC-BY</td>
<td>EN</td>
<td>Repository Issues Requirements Specification Shapes</td>
<td>Core ontology that defines the necessary resources to create a mapping.</td>
</tr>
<tr>
<td>RML-ID: Source and Target</td>
<td>OWL</td>
<td>CC-BY</td>
<td>EN</td>
<td>Repository Issues Requirements Specification Shapes</td>
<td>Ontology module that allows the description of input data sources and target outputs.</td>
</tr>
<tr>
<td>RML-CC: Collections and Containers</td>
<td>OWL</td>
<td>CC-BY</td>
<td>EN</td>
<td>Repository Issues Requirements Specification Shapes</td>
<td>Ontology module that allows the generation of collections and containers.</td>
</tr>
<tr>
<td>RML-FNML: Functions</td>
<td>OWL</td>
<td>CC-BY</td>
<td>EN</td>
<td>Repository Issues Requirements Specification Shapes</td>
<td>Ontology module that allows the application of data transformation functions.</td>
</tr>
<tr>
<td>RML-Star</td>
<td>OWL</td>
<td>CC-BY</td>
<td>EN</td>
<td>Repository Issues Requirements Specification Shapes</td>
<td>Ontology module that allows the construction of RDF-star graphs.</td>
</tr>
</tbody>
</table>

http://w3id.org/rml/portal
Test cases (January - March 2024)

Track 1: Conformance

The set of new specification for the RDF Mapping Language (RML) established by the W3C Community Group on Knowledge Graph Construction provide a set of test-cases for each module:

- RML-Core
- RML-IO
- RML-CC
- RML-FNML
- RML-Star

These test-cases are evaluated in this Track of the Challenge to determine their feasibility, correctness, etc. by applying them in implementations. **This Track is in Beta status because these new specifications have not seen any implementation yet, thus it may contain bugs and issues.** If you find problems with the mappings, output, etc. please report them to the corresponding repository of each module.

**Note:** validating the output of the RML Star module automatically through the provided tooling is currently not possible, see [https://github.com/kg-construct/challenge-tool/issues/1](https://github.com/kg-construct/challenge-tool/issues/1).

Through this Track we aim to spark development of implementations for the new specifications and improve the test-cases. Let us know your problems with the test-cases and we will try to find a solution.
Paper presented at ISWC

Best Paper Candidates

- **Session 1A (Wednesday):** Michael Färber, Johan Krause, David Lamprecht, Linn Aung and Peter Haase. *SemOpenAlex: The Scientific Landscape in 26 Billion RDF Triples*

- **Session 10A (Friday):** Ana Iglesias-Molina, Dylan Van Assche, Julián Arenas-Guerrero, Ben De Meester, Christophe Debruyne, Samaneh Jozashoori, Pano Maria, Franck Michel, David Chaves-Fraga and Anastasia Dimou. *The RML Ontology: A Community-Driven Modular Redesign After a Decade of Experience in Mapping Heterogeneous Data to RDF*

- **Session 10A (Friday):** Jacopo de Berardinis, Valentina Anita Carriero, Nitisha Jain, Nicolas Lazzari, Albert Moreno-Peñuela, Andrea Poltronieri and Valentina Presutti. *The Polifonia Ontology Network: Building a Semantic Backbone for Musical Heritage*
Meeting on Santiago de Compostela - December 2023

- 2 days workshop
- Discussion on open issues
- Agreement on the transition from the CG to a W3C WG
- RML-joins transformed into RML-LogicalViews

Supported by CA Distributed KG
…but we are not done! (although we are almost there)

Next steps:

- Feedback on current work
- Opinions about current open issues
- **Started the transition into a W3C Working Group (end of 2024)**
- **Engines already implemented the new specs!** (Wait for the challenge results)

Join us!

- [public-kg-construct@w3.org](mailto:public-kg-construct@w3.org)
- [w3id.org/kg-construct](http://w3id.org/kg-construct)
- [kg-construct.slack.com](http://kg-construct.slack.com)
W3C Community Group
Knowledge Graph Construction

http://w3id.org/kg-construct