

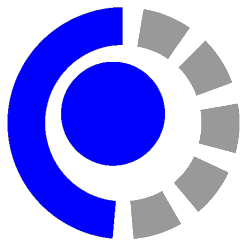
Agentes Inteligentes Para La Web

R2RML: RDB To RDF Mapping Language

Manuel Latorre FAI-1931
manuel.latorre@est.fi.uncoma.edu.ar

Sergio Strazzacappa FAEA-1432
sergio.strazzacappa@est.fi.uncoma.edu.ar

Primer cuatrimestre 2023



Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE



Índice

1. Introducción	1
2. Marco Teórico	1
3. Ejemplo	2

Índice de figuras

1. Ejemplo de un archivo de mapeo	3
2. Ejemplo de mapeo de una relacion muchos a muchos	3
3. Grafo de conocimiento obtenido a partir del mapeo	4
4. Consulta de los departamentos de cada empleado	4

1. Introducción

El objetivo de **R2RML**(RDB to RDF mapping language) es especificar transformaciones de información disponibles de bases de datos relacionales en datasets RDF materializados o virtuales. Convertir el contenido de las bases de datos relacionales a RDF juega un rol fundamental en la *Web of Data* debido a que esta utiliza RDF para representar y enlazar datos de forma estructurada.

Al convertir los datos relacionales en RDF, se puede aprovechar el poder de consultas y análisis avanzados en la Web Semántica. Esto se debe a que RDF proporciona un modelo flexible y expresivo para representar los datos lo que permite un análisis mas profundo que incluye inferencias y razonamiento.

Además esto permite la construcción de grafos de conocimiento (KG o *knowledge graph*) que es una de las tecnologías mas utilizadas para la integración de datos. Para la construcción de estos KG desde una fuente de datos no RDF, el mapeo de lenguajes permite definir relaciones entre fuentes de datos de entrada y ontologías de una forma declarativa y mantenible

2. Marco Teórico

R2RML es una recomendación W3C que permite expresar mapeos personalizados desde bases de datos relacionales a RDF. Un documento de mapeo R2RML esta compuesto por un conjunto de tripletas de la forma `rr:TriplesMap`, donde *rr* es el prefijo de la siguiente URI:

<http://www.w3.org/ns/r2rml#>.

Estas tripletas consisten de tres propiedades:

1. `rr:logicalMap` especifica la tabla logica a ser mapeada, puede ser una tabla SQL, una vista SQL o una vista R2RML
2. `rr:subjectMap` especifica la clase objetivo y la forma de generación del URI
3. `rr:predicateObjectMap` especifica la propiedad objetivo y la generación del objeto via `rr:objectMap`, cuyo valor puede ser obtenido por constantes (`rr:constant`), columnas (`rr:column`) o template (`rr:template`). Las vinculaciones con otras tablas se definen con el uso de `rr:refObjectMap`, que especifica las tripletas a mapear y las condiciones de unión

Algunos posibles usos de R2RML pueden ser:

- Integrar datos almacenados en bases de datos relacionales con el ecosistema semántico de la Web. Permite convertir datos estructurados en RDF, lo que facilita su interoperabilidad y enriquecimiento mediante el uso de ontologías y recursos enlazados. Por ejemplo en entornos empresariales, es común tener múltiples sistemas de bases de datos relacionales que almacenan datos en diferentes formatos y esquemas. Utilizando R2RML, se pueden crear mapeos entre estos sistemas y un modelo de datos RDF común, lo que facilita la integración y la realización de consultas unificadas en los datos empresariales.
- Permite publicar datos de bases de datos relacionales en la Web Semántica, lo que facilita su acceso y reutilización por parte de aplicaciones y sistemas que consumen datos en formato

RDF. Esto ayuda a mejorar la visibilidad y accesibilidad de los datos, promoviendo su uso en diferentes dominios. Los gobiernos y las instituciones gubernamentales están adoptando cada vez más la publicación de datos abiertos en la Web Semántica. R2RML se utiliza para convertir datos de bases de datos gubernamentales en formato RDF, lo que permite una mayor interoperabilidad y reutilización de los datos por parte de ciudadanos, investigadores y aplicaciones gubernamentales.

- En entornos donde se utilizan diferentes bases de datos relacionales con esquemas y estructuras diferentes, R2RML proporciona una solución para integrar estas fuentes de datos heterogéneas. Permite definir mapeos personalizados entre los esquemas de las bases de datos y el modelo RDF, lo que facilita la unificación y consulta de datos desde una perspectiva semántica. En el campo de la salud, existen múltiples sistemas de bases de datos relacionales que almacenan información sobre pacientes, historias clínicas, medicamentos, etc. Utilizando R2RML, es posible integrar estos datos heterogéneos en un modelo de datos RDF común, lo que facilita la interoperabilidad entre diferentes sistemas de información de salud.

3. Ejemplo

Como se puede observar en la imagen 1 se esta haciendo un mapeo a partir de una tabla logica llamada DEPT (la cual se encuentra cargada en un archivo .csv). Luego se definen los predicados de estas referenciando a las columnas de dicha tabla. Como resultado se obtendra un mapeo con tripletas de la forma:

Tripletas obtenidas en el mapeo

```
<http://data.example.com/department/10>
  <http://example.com/ns#dname> "APPSERVER" .
<http://data.example.com/department/10>
  <http://example.com/ns#location> "NEW_YORK" .

<http://data.example.com/department/20>
  <http://example.com/ns#dname> "RESEARCH" .
<http://data.example.com/department/20>
  <http://example.com/ns#location> "BOSTON" .
```

Por otro lado para definir una relación, por ejemplo muchos a muchos, se puede utilizar `rr:template` para referenciar a la URI de las tripletas generadas por otros mapeos como se puede ver en la figura 2.

Finalmente se generara un grafo de conocimiento representado en RDF. El ejemplo mostrado son solo partes de un mapeo mucho mas grande, si el resultado de dicho mapeo se carga en la plataforma Metaphactory se logra visualizar el grafo de conocimiento de la figura 3. Además se pueden definir consultas en SPARQL como se hace en la figura 4

```
1 @prefix rr: <http://www.w3.org/ns/r2rml#>.
2 @prefix ex: <http://example.com/ns#>.
3
4 <#TriplesMap2>
5   rr:logicalTable [ rr:tableName "DEPT" ];
6   rr:subjectMap [
7     rr:template "http://data.example.com/department/{DEPTNO}";
8     rr:class ex:Department;
9   ];
10  rr:predicateObjectMap [
11    rr:predicate ex:dname;
12    rr:objectMap [ rr:column "DNAME" ];
13  ];
14  rr:predicateObjectMap [
15    rr:predicate ex:location;
16    rr:objectMap [ rr:column "LOC" ];
17  ].
```

Figura 1: Ejemplo de un archivo de mapeo

```
1 @prefix rr: <http://www.w3.org/ns/r2rml#>.      Wishmaster, 49 minutes ago • Agreg
2 @prefix ex: <http://example.com/ns#>.
3
4 <#TriplesMap3>
5   rr:logicalTable [ rr:tableName "EMP2DEPT" ];
6   rr:subjectMap [
7     rr:template "http://data.example.com/employee/{EMPNO}";
8   ];
9   rr:predicateObjectMap [
10    rr:predicate ex:department;
11    rr:objectMap [ rr:template "http://data.example.com/department/{DEPTNO}" ];
12  ].
```

Figura 2: Ejemplo de mapeo de una relacion muchos a muchos

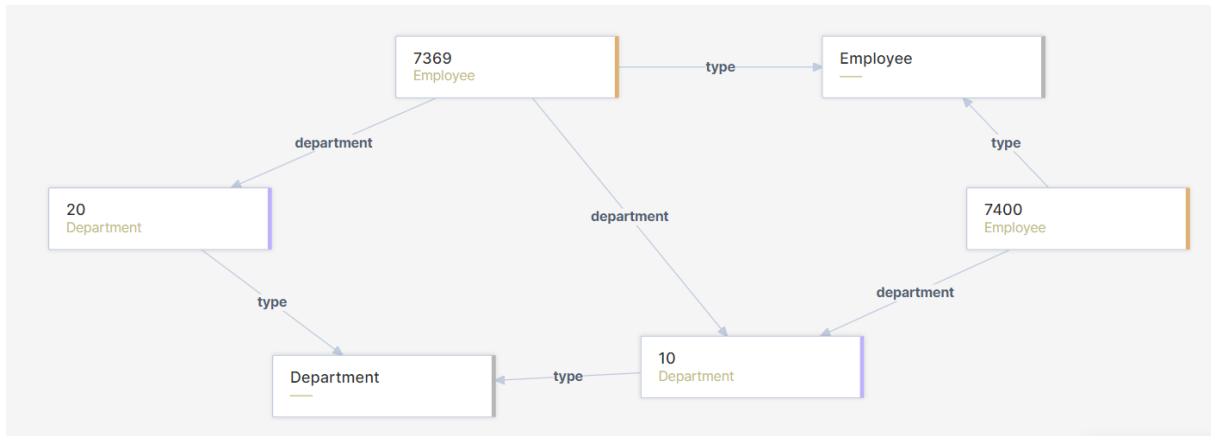


Figura 3: Grafo de conocimiento obtenido a partir del mapeo

```

1 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX emp: <http://data.example.com/employee/>
4 PREFIX pred: <http://example.com/ns#>
5 SELECT * WHERE {
6   ?emp pred:department ?departamentos
7 } LIMIT 10

```

Repository: (from context) **Execute** **Save** Query Execution Time: 199 ms / 0.20 s

Table **Fetch Labels: OFF**

Quick search

3 entries

emp	departamentos
http://data.example.com/employee/7369	http://data.example.com/department/10
http://data.example.com/employee/7369	http://data.example.com/department/20
http://data.example.com/employee/7400	http://data.example.com/department/10

1 Show 10 rows

Figura 4: Consulta de los departamentos de cada empleado

Referencias

- [1] Freddy Priyatna, Oscar Corcho y Juan Sequeda. «Formalisation and experiences of R2RML-based SPARQL to SQL query translation using morph». En: *Proceedings of the 23rd International Conference on World Wide Web*. 2014, págs. 479-490.
- [2] Julián Arenas-Guerrero et al. «Knowledge graph construction with R2RML and RML: an ETL system-based overview». En: *CEUR workshop proceedings*. Vol. 2873. CEUR Workshop Proceedings. 2021.
- [3] Richard Cyganiak Souripriya Das Seema Sundara. «R2RML: RDB to RDF Mapping Language». En: *R2RML: RDB to RDF Mapping Language*. W3C. 2012. URL: <https://www.w3.org/TR/r2rml/>.