

D2RQ Mapper

Yasunori Yamamoto and Toshiaki Katayama

Database Center for Life Science, Univ. of Tokyo Kashiwa-no-ha Campus Station Satellite 6F.
178-4-4 Wakashiba, Kashiwa-shi, Chiba, JAPAN
`{yy,ktym}@dbcls.rois.ac.jp`

Abstract. D2RQ Mapper (<http://d2rq.dbcls.jp/>) is a web application to edit a mapping file of D2RQ, a middleware to bridge Relational Database (RDB) and Resource Description Framework (RDF). A D2RQ mapping file defines how to map data stored in an RDB to RDF in the turtle format, and to write it by a text editor is cumbersome. D2RQ Mapper assists you to edit it by contextualizing input forms in the mapping language. In addition, D2RQ Mapper supports the R2RML format to output a mapping definition. We provide a Docker image of D2RQ Mapper, and so you can easily use it within your intranet.

Keywords. D2RQ, R2RML, SPARQL, SQL, RDF, RDB

1 Introduction

To facilitate publishing datasets accessible with the Semantic Web technology such as SPARQL¹ or Linked (Open) Data², an easy way of transition from datasets stored in relational databases (RDBs) to those in Resource Description Framework (RDF) is crucial. However, it is time-consuming for many RDB maintainers to replicate their entire databases into RDF. To alleviate this task and fully utilize data stored in RDB in the Semantic Web context while keeping the database content up-to-date, it is ideal to have a tool to map RDB datasets to those of RDF graph and to provide an interface where user can access to the RDB using SPARQL.

One of the widely used such tools is D2RQ³, which provides a SPARQL endpoint that connects to a designated RDB SQL interface and also has a capability of dumping RDB datasets in RDF. D2RQ needs a configuration or mapping file that describes how to map an RDB schema to a Semantic Web model in the D2RQ mapping language. Although D2RQ assists the transition from RDB to RDF, to write a mapping file is not straightforward unless users satisfy one created automatically by a D2RQ tool. D2RQ Mapper is a web application that assists generating a mapping file for the D2RQ platform. Users can create a mapping file without a text editor and decrease syntactic errors. D2RQ Mapper shows a schema of a given database and valid options conforming to the mapping language. Although users can only use a

¹ <http://www.w3.org/TR/sparql11-overview/>

² <http://www.w3.org/DesignIssues/LinkedData.html>

³ <http://d2rq.org/>

subset of the specifications that the language provides by D2RQ Mapper currently, based on our preliminary survey, we believe it helps to create a configuration file a lot, and we continue to improve to support more specifications.

With the D2RQ Mapper, user can easily

1. create a configuration file in D2RQ Mapping Language or R2RML,
2. query the RDB databases with SPARQL on-the-fly via a D2RQ server, and,
3. dump a resulting RDF dataset.

The advantage of using our tool is that you can easily edit and try a configuration file within the tool without rebooting a D2RQ server. We focus on alleviating any troubles to generate an ideal mapping file, and usually iteration of edit and try is needed based on our mapping experience.

D2RQ Mapper is available on this web site (<http://d2rq.dbcls.jp/>) for externally accessible RDBs, and its Docker image (<https://hub.docker.com/r/d2rqmapper/d2rq-mapper/>) is available for locally accessible RDBs within an intranet.

2 Implementation

As mentioned above, we provide D2RQ Mapper in two ways, that is, a globally accessible web application and a Docker image. Once the latter is deployed, there is no difference between them except for the sign-in methods. With the former, a user can use D2RQ Mapper by Twitter or Facebook accounts in addition to one created at the D2RQ Mapper site. On the other hand, a user can only use an account created at the deployed D2RQ Mapper in the case of the latter. This is because to use D2RQ Mapper within an intranet generally needs a basic user identification function only. The way of deploying from the Docker image in a Docker terminal is as follows.

```
$ docker pull d2rqmapper/d2rq-mapper
$ docker run -d -p 80:80 d2rqmapper/d2rq-mapper
```

Then, you can use D2RQ Mapper from your browser by accessing to the IP address shown when launching the Docker terminal.

D2RQ Mapper runs on Ruby on Rails 4.2 with PostgreSQL 9.3. The version of D2RQ tools used within D2RQ Mapper is 0.8.3dev obtained from the GitHub⁴. Data concerning users, designated RDB Management Systems, and mapping to RDF models are stored in the PostgreSQL. A D2RQ mapping file and an R2RML mapping file are generated from the stored data with D2RQ Mapper, which does not use the generate-mapping tool of the D2RQ package. An RDF dataset is generated using the dump-rdf tool of the D2RQ package, and a given SPARQL query is processed using the d2r-query, which is also included in the package.

We designed D2RQ Mapper to minimize simple syntactic errors and improve efficiency to edit a mapping file. This policy is reflected in the way of navigating the users. After signing-in D2RQ Mapper, a user needs to configure an RDB setting along with a name of the mapping to be generated first. When proceeding to the next step,

⁴ <https://github.com/d2rq/d2rq>

D2RQ Mapper verifies it and retrieves the schema of the given RDB such as the data of its tables and their columns. Therefore, D2RQ Mapper can show the user the table list at the following step, and the user can determine which tables to be included to generate a target RDF dataset. In addition, a user can set up a join table that can be seen as an ordinary table by defining how to join two or three tables. In the same way, a user can configure how each column value is mapped to the target RDF dataset in each table. On the configuration page, D2RQ Mapper shows the user sample row data of the table to facilitate referring to the actual data in the RDB.

Once configuring the RDB setting, a user can get the resultant mapping file at any-time. In addition, the target RDF dataset is downloadable. We expect that D2RQ Mapper accelerate the transition from RDB to RDF.

3 Future works

The current version of D2RQ Mapper does not accept an ontology or a vocabulary. To enable a user to edit a mapping file by referring to it within D2RQ Mapper, we are adding a function of providing a graphical interface that shows an ontology map. In addition, we will improve D2RQ Mapper to support more specifications of D2RQ Mapping Language and R2RML.

Acknowledgments. This work was supported by the National Bioscience Database Center (NBDC) of the Japan Science and Technology Agency (JST).