Graphical display of SPARQL queries with arq, rapper and graphviz

mja@statgroup.dk

2016-06-29

Contents

Graphical display of results from SPARQL scripts for the demographics cube (DC-DEMO-sample.ttl)

The examples below uses `arq` from Apache Jena (http://jena.apache.org). To install arq - download and unpack the latest version of apache-jena from (http://jena.apache.org/download/index.cgi). Put the executable `arq` in the path, or invoke `arq` with the full path to the directory with arq.

The use of arq is described many places, see for example (http://www.learningsparql.com/).

All `arq` commands below are to be run in the directory with the sample files, which is `inst/extdata/sample-rdf` directory or `extdata/sample-rdf` depending on the whether the development version or the installed version of the package is used.

The cd below in each code block is included because I could not find a quick way to get the code chunk executed in that directory. knitr is flexible enough to do it, I have not yet found the right way to do it. So, ignore the repeated cd ..

Get one observation and display graphically

The SPARQL query returns all triples for one observation, and stores it in the file `fordot.ttl`. The file is used as input to `rapper` and converted to the dot format, and displayed using the `dot` program part of Graphviz (http://www.graphviz.org/).

```
cd ../extdata/sample-rdf
arq --data DC-DEMO-sample.ttl --query OneQBobservation.rq > fordot.ttl
cat fordot.ttl
rapper -i turtle -o dot fordot.ttl > fordot.dot
dot -x -Tpdf -ograph.pdf fordot.dot
```

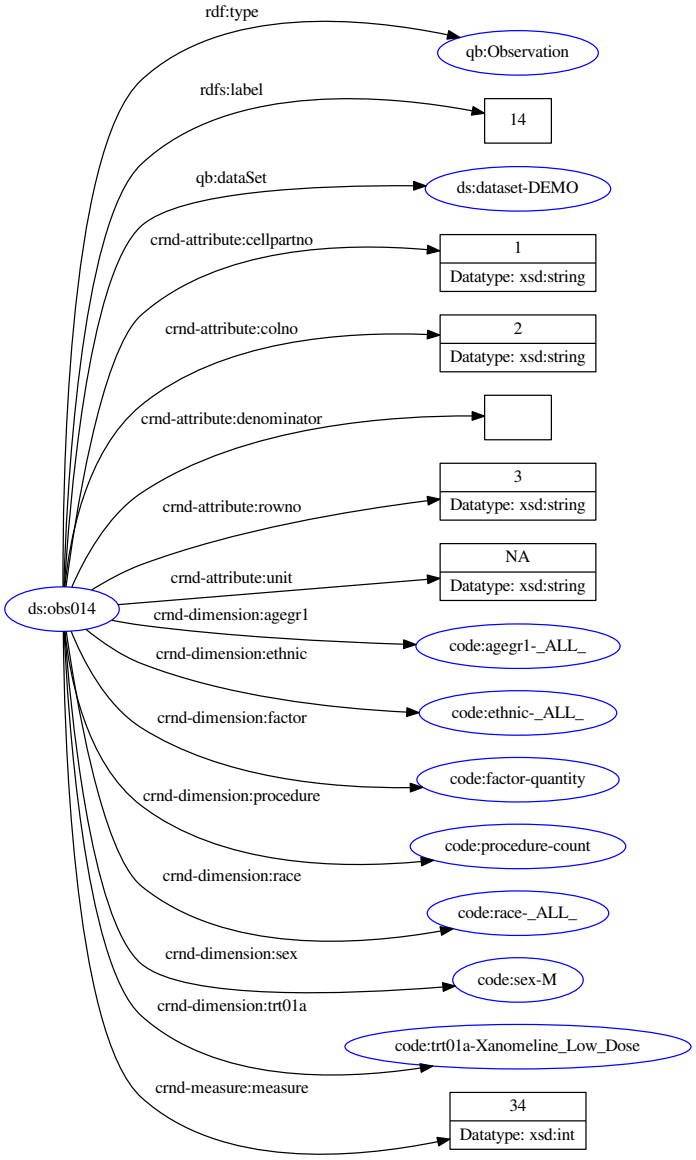
```
## @prefix dccs:  <http://www.example.org/dc/demo/dccs/> .
## @prefix code:  <http://www.example.org/dc/code/> .
## @prefix sdtms-1-3: <http://rdf.cdisc.org/sdtm-1-3/schema#> .
## @prefix adam-2-1: <http://rdf.cdisc.org/std/adam-2-1#> .
## @prefix owl:   <http://www.w3.org/2002/07/owl#> .
## @prefix xsd:   <http://www.w3.org/2001/XMLSchema#> .
## @prefix sdtm-1-3: <http://rdf.cdisc.org/std/sdtm-1-3#> .
```

```
## @prefix cdash-1-1: <http://rdf.cdisc.org/std/cdash-1-1#> .
## @prefix skos:  <http://www.w3.org/2004/02/skos/core#> .
## @prefix rdfs:  <http://www.w3.org/2000/01/rdf-schema#> .
## @prefix adamvr-1-2: <http://rdf.cdisc.org/std/adamvr-1-2#> .
## @prefix crnd-attribute: <http://www.example.org/dc/attribute#> .
## @prefix sdtm-1-2: <http://rdf.cdisc.org/std/sdtm-1-2#> .
## @prefix ds:    <http://www.example.org/dc/demo/ds/> .
## @prefix sdtmct: <http://rdf.cdisc.org/sdtm-terminology#> .
## @prefix qb:    <http://purl.org/linked-data/cube#> .
## @prefix mms:   <http://rdf.cdisc.org/mms#> .
## @prefix crnd-dimension: <http://www.example.org/dc/dimension#> .
## @prefix dct:   <http://purl.org/dc/terms/> .
## @prefix cdiscs: <http://rdf.cdisc.org/std/schema#> .
## @prefix cdashct: <http://rdf.cdisc.org/cdash-terminology#> .
## @prefix dcat:  <http://www.w3.org/ns/dcat#> .
## @prefix prov:  <http://www.w3.org/ns/prov#> .
## @prefix sdtmig-3-1-3: <http://rdf.cdisc.org/std/sdtmig-3-1-3#> .
## @prefix crnd-measure: <http://www.example.org/dc/measure#> .
## @prefix adamig-1-0: <http://rdf.cdisc.org/std/adamig-1-0#> .
## @prefix cts:   <http://rdf.cdisc.org/ct/schema#> .
## @prefix pav:   <http://purl.org/pav> .
## @prefix sdtmig-3-1-2: <http://rdf.cdisc.org/std/sdtmig-3-1-2#> .
## @prefix sendig-3-0: <http://rdf.cdisc.org/std/sendig-3-0#> .
## @prefix adamct: <http://rdf.cdisc.org/adam-terminology#> .
## @prefix rdf:   <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
## @prefix sendct: <http://rdf.cdisc.org/send-terminology#> .
## @prefix rrdfqbcrnd0: <http://www.example.org/rrdfqbcrnd0/> .
## @prefix dc:    <http://purl.org/dc/elements/1.1/> .
##
## ds:obs014  a                        qb:Observation ;
##         rdfs:comment                "Statistic for number of records/Statistics for factor with the
##         rdfs:label                  "14" ;
##         qb:dataSet                  ds:dataset-DEMO ;
##         crnd-attribute:cellpartno   "1" ;
##         crnd-attribute:colno        "2" ;
##         crnd-attribute:denominator  "" ;
##         crnd-attribute:measurefmt   "%6.0f" ;
##         crnd-attribute:rowno        "3" ;
##         crnd-attribute:unit         "NA" ;
##         crnd-dimension:agegr1       code:agegr1-_ALL_ ;
##         crnd-dimension:ethnic       code:ethnic-_ALL_ ;
##         crnd-dimension:factor       code:factor-quantity ;
##         crnd-dimension:procedure    code:procedure-count ;
##         crnd-dimension:race         code:race-_ALL_ ;
##         crnd-dimension:sex          code:sex-M ;
##         crnd-dimension:trt01a       code:trt01a-Xanomeline_Low_Dose ;
##         crnd-measure:measure        "34"^^xsd:double .
## rapper: Parsing URI file:///home/ma/projects/rrdfqbcrnd/rrdfqbcrndex/inst/extdata/sample-rdf/fordot.
## rapper: Serializing with serializer dot
## rapper: Parsing returned 18 triples
```

The pdf file can then be viewed using a pdf viewer.

Alternatively, knitr supports presenting dot as show below.

```
digraph {
    rankdir = LR;
    charset="utf-8";

    "Rds:obs014" -> "Rqb:Observation" [ label="rdf:type" ];
    "Rds:obs014" -> "L14" [ label="rdfs:label" ];
    "Rds:obs014" -> "Rds:dataset-DEMO" [ label="qb:dataSet" ];
    "Rds:obs014" -> "L1|Datatype: xsd:string" [ label="crnd-attribute:cellpartno" ];
    "Rds:obs014" -> "L2|Datatype: xsd:string" [ label="crnd-attribute:colno" ];
    "Rds:obs014" -> "L" [ label="crnd-attribute:denominator" ];
    "Rds:obs014" -> "L3|Datatype: xsd:string" [ label="crnd-attribute:rowno" ];
    "Rds:obs014" -> "LNA|Datatype: xsd:string" [ label="crnd-attribute:unit" ];
    "Rds:obs014" -> "Rcode:agegr1-_ALL_" [ label="crnd-dimension:agegr1" ];
    "Rds:obs014" -> "Rcode:ethnic-_ALL_" [ label="crnd-dimension:ethnic" ];
    "Rds:obs014" -> "Rcode:factor-quantity" [ label="crnd-dimension:factor" ];
    "Rds:obs014" -> "Rcode:procedure-count" [ label="crnd-dimension:procedure" ];
    "Rds:obs014" -> "Rcode:race-_ALL_" [ label="crnd-dimension:race" ];
    "Rds:obs014" -> "Rcode:sex-M" [ label="crnd-dimension:sex" ];
    "Rds:obs014" -> "Rcode:trt01a-Xanomeline_Low_Dose" [ label="crnd-dimension:trt01a" ];
    "Rds:obs014" -> "L34|Datatype: xsd:int" [ label="crnd-measure:measure" ];

    // Resources
    "Rds:obs014" [ label="ds:obs014", shape = ellipse, color = blue ];
    "Rqb:Observation" [ label="qb:Observation", shape = ellipse, color = blue ];
    "Rds:dataset-DEMO" [ label="ds:dataset-DEMO", shape = ellipse, color = blue ];
    "Rcode:agegr1-_ALL_" [ label="code:agegr1-_ALL_", shape = ellipse, color = blue ];
    "Rcode:ethnic-_ALL_" [ label="code:ethnic-_ALL_", shape = ellipse, color = blue ];
    "Rcode:factor-quantity" [ label="code:factor-quantity", shape = ellipse, color = blue ];
    "Rcode:procedure-count" [ label="code:procedure-count", shape = ellipse, color = blue ];
    "Rcode:race-_ALL_" [ label="code:race-_ALL_", shape = ellipse, color = blue ];
    "Rcode:sex-M" [ label="code:sex-M", shape = ellipse, color = blue ];
    "Rcode:trt01a-Xanomeline_Low_Dose" [ label="code:trt01a-Xanomeline_Low_Dose", shape = ellipse, colo

    // Anonymous nodes

    // Literals
    "L14" [ label="14", shape = record ];
    "L1|Datatype: xsd:string" [ label="1|Datatype: xsd:string", shape = record ];
    "L2|Datatype: xsd:string" [ label="2|Datatype: xsd:string", shape = record ];
    "L" [ label="", shape = record ];
    "L3|Datatype: xsd:string" [ label="3|Datatype: xsd:string", shape = record ];
    "LNA|Datatype: xsd:string" [ label="NA|Datatype: xsd:string", shape = record ];
    "L34|Datatype: xsd:int" [ label="34|Datatype: xsd:int", shape = record ];

    label="\n\nModel:\n(Unknown)\n\nNamespaces:\nprov: http://www.w3.org/ns/prov#\ncrnd-dimension: http
}
```



Model:
(Unknown)

Namespaces:
prov: http://www.w3.org/ns/prov#
crnd-dimension: http://www.example.org/dc/dimension#
mms: http://rdf.cdisc.org/mms#
crnd-measure: http://www.example.org/dc/measure#
code: http://www.example.org/dc/code/
qb: http://purl.org/linked-data/cube#
dccs: http://www.example.org/dc/demo/dccs/
rdfs: http://www.w3.org/2000/01/rdf-schema#
dcat: http://www.w3.org/ns/dcat#
pav: http://purl.org/pav
dct: http://purl.org/dc/terms/
xsd: http://www.w3.org/2001/XMLSchema#
owl: http://www.w3.org/2002/07/owl#
rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
skos: http://www.w3.org/2004/02/skos/core#
cts: http://rdf.cdisc.org/ct/schema#
rrdfqbcrnd0: http://www.example.org/rrdfqbcrnd0/
ds: http://www.example.org/dc/demo/ds/
crnd-attribute: http://www.example.org/dc/attribute#

Figure 1: SPARQL query