Execute SPARQL integrity constraints on a data cube

mja@statgroup.dk

2015-01-11

Contents

```
library(rrdf)
```

```
## Loading required package: rJava
## Loading required package: methods
## Loading required package: rrdflibs
```

```
library(rrdfqbcrnd0)
```

```
## Loading required package: xlsx
## Loading required package: xlsxjars
## Loading required package: RCurl
## Loading required package: bitops
##
## Attaching package: 'RCurl'
##
## The following object is masked from 'package:rJava':
##
##     clone
```

```
## Warning: replacing previous import by 'rJava::clone' when loading
## 'rrdfqbcrnd0'
```

```
cubeData = new.rdf(ontology=FALSE)

myprefixes= c(Get.default.crnd.prefixes(), list(
 "dccs"="http://www.example.org/dc/demog/dccs/" ,
 "code"="http://www.example.org/dc/code/" ,
 "ds"=  "http://www.example.org/dc/demog/dataset/" ,
 "prop"="http://www.example.org/dc/demog/prop/",
 "valid"="http://www.example.org/dc/demog/valid/"
  ))

forsparqlprefix= paste("prefix", paste(names(myprefixes),":",sep=""), paste("<",myprefixes,">",sep=""),

qbfile<- system.file("extdata/sample-rdf", "DC-DM-sample.TTL", package="rrdfqbcrnd0")
load.rdf( qbfile, format="N3", appendTo= cubeData)

ic1<- "ASK {
  {
    # Check observation has a data set
    ?obs a qb:Observation .
```

```
    FILTER NOT EXISTS { ?obs qb:dataSet ?dataset1 . }
  } UNION {
    # Check has just one data set
    ?obs a qb:Observation ;
       qb:dataSet ?dataset1, ?dataset2 .
    FILTER (?dataset1 != ?dataset2)
  }
}"
```

rrdf does not implement the ASK query. As a word around use transform to, most likely more inefficient, a `SELECT (COUNT(*) ...)` query.

```
ic1Select<- gsub("ASK \\{", "SELECT (COUNT(*) AS ?nfail) WHERE \\{", ic1)

cube.ic1= sparql.rdf(cubeData, paste( forsparqlprefix, ic1Select  )  )
print(cube.ic1)
```

 nfail

[1,] "0"