

Integrációs és ellenőrzési technikák

1. gyakorlat

SPARQL lekérdezések használata RDF adatbázison

Az alábbi feladatokat a virtuális gépre telepített RDF4J adatbázis workbench felületén oldja meg:

<http://localhost:8080/rdf4j-workbench/>

Az OpenRDF Workbench felület angol nyelvű kézikönyvét itt találja:

http://docs.rdf4j.org/server-workbench-console/#_rdf4j_workbench

A SPARQL lekérdező nyelv specifikációja:

<https://www.w3.org/TR/sparql11-query/>

A felhasznált adatbázis a Szépművészeti Múzeumban tárolt műkincsekről tartalmaz információt.

<https://datahub.io/dataset/data-szepmuveszeti-hu>

0) Töltse be az OpenRDF Workbench-be a minta adatbázist.

Nyissa meg böngészőben az OpenRDF Workbench felületét:

<http://localhost:8080/rdf4j-workbench/>

A bal oldali menüben válassza a “Repositories / New repository” menüpontot, majd az alábbi beállításokkal hozhatja létre az adatbázist:

- Type: Native Java Store
(“In Memory Store” esetén előfordulhat adatvesztés az adatbázis leállításakor.)
- ID: szepmuveszeti
- Next
- Triple indexes: spoc, posc, csop, csop, cops
Az alapértelmezett (spoc, posc) helyett további indexek megadásával a lekérdezések futtatása nagy adatbázisban gyorsabb lehet.
- Create

A fejlécben az “Repository:” után látja mindig az aktuálisan kiválasztott adatbázis azonosítóját. Ha több adatbázisa is van, váltani a “Repositories” menüpontban tud.

Az adatok betöltéséhez a “Modify / Add” menüpontban:

- Kapcsolja ki a “use base URI as context identifier” checkbox-ot.
- RDF Data File: Choose File
Válassza ki az asztalon a “szepmuveszeti.20151028.n t” file-t.
- Data format: N3
Az “(autodetect)” erre a file-ra nem működik. A file tallózása után állítsa be, különben visszaállítódik.
- Upload

Ha később újra akarja kezdeni a munkát, a “Repositories / Delete repository” menüpontban törölheti az adatbázist, és újra létrehozhatja.

1) Keresse meg az adatbázisban az alkotókat. Milyen adatok szerepelnek róluk?

Az OpenRDF Workbench-ben a "Types" fül tartalmazza az adatbázisban előforduló típusokat. Az egyes típusra kattintva minden olyan RDF állítást megkapunk, amiben a kiválasztott típus szerepel.

RDF-ben egy entitás típusát a “?s rdf:type ?o” állítás fejezi ki, ahol a subject helyén (?s) áll a tipizált entitás, és az object (?o) helyén a típus.

A Workbench felületen a listázott állítások táblázatában az azonosítókra kattintva láthatunk minden olyan állítást, amiben az adott entitás szerepel. Így bejárhatjuk az adatbázisban szereplő kapcsolatokat.

Keresse meg azt a típust, amihez az alkotók tartoznak. Milyen adatokat tartalmaz az adatbázis az alkotókról?

2) Az adatbázisban elválik a "műalkotás" és az "alkotási folyamat". Mik ezek a típusok, és hogyan kapcsolódnak egymáshoz?

A feladatot egy-egy alkotóból kiindulva, a gráf bejárásával oldja meg.

3) Hogyan tárolja az adatbázis a műalkotások méretét? Miért?

4) Készítsen SPARQL lekérdezést, ami megtalálja az összes alkotót.

A Workbench felületen az “Explore / Query” funkcióval tud futtatni SPARQL lekérdezéseket. A lekérdezések elején meg kell adnia a használt prefix-eket, amiket a felület az adatbázis tállózásánál és a lekérdezések eredményeinek megjelenítésénél figyelembe vesz. Az adatbázisban használt prefixeket az “Explore / Namespaces” oldalon tudja megnézni. Az “rdf” beépített prefix, ezért nem kell megadni.

Például:

```
PREFIX ecrm: <http://erlangen-crm.org/current/>
SELECT ?s ?v {
    ?s rdf:type ecrm:E54_Dimension ;
        ecrm:P90_has_value ?v
}
```

Az “rdf:type” szinonímájaként használhatja az “a” kulcsszót.

5) Készítsen SPARQL lekérdezést, ami visszaadja a "Giovanni" nevű alkotókat (akiknek a nevében szerepel a "Giovanni" string). A kapott táblázat két oszlopa legyen a művész azonosítója és a neve.

A névben történő kereséshez használja a "regex" függvényt:

<https://www.w3.org/TR/sparql11-query/#func-regex>

```
FILTER regex(?name, "giovanni", "i")
```

6) Kapcsolja össze a művészeket az alkotásaikkal. A 2)-es feladat eredményeit kell használnia.

A megoldás oszlopai: alkotó, alkotási folyamat, műalkotás.

Házi feladat

Rajzolja le az adatbázis sémáját gráfként.

A lényeges típusok legyenek a gráf csomópontjai, és a köztük előforduló tulajdonságok a gráf élei.

Írjon lekérdezést, ami megtalálja azokat a "rézkarc" technikával készült műveket, melyeket Rembrandt készített.