



C3V9

Návrh zajištění podpory asistenta pro dotazování nad 5* datovými sadami

Vytvořeno v rámci projektu

Rozvoj datových politik v oblasti zlepšování kvality
a interoperability dat veřejné správy
CZ.03.4.74/0.0/0.0/15_025/0013983

Klíčová aktivita: 04: Návrhy a realizace opatření pro zvyšování povědomí
o otevřených datech

Indikátor: 8 05 00 Počet napsaných a zveřejněných analytických a strategických
dokumentů (vč. evaluačních)

Verze výstupu: 01



Podpora asistenta dotazování nad 5* datovými sadami

Rozšíření Národního katalogu otevřených dat o podporu datových služeb a indikaci jejich specifikací, které bylo navrženo ve výstupu C2V2 a Ministerstvem vnitra už i implementováno do Národního katalogu otevřených dat (NKOD) umožnilo rozpoznání datových sad publikovaných v 5* pomocí SPARQL endpointu. Pro takové distribuce datových sad je nyní automaticky vygenerován odkaz na YASGUI, asistenta dotazování v jazyce SPARQL. Asistent je automaticky nasměrován na SPARQL endpoint dané datové sady, automaticky je také předvyplněn a spuštěn iniciální SPARQL dotaz. Jak URL instance YASGUI, tak výchozí dotaz jsou konfigurovatelné.

Integrace do Národního katalogu otevřených dat

Do uživatelského rozhraní Národního katalogu otevřených dat je tedy přidán odkaz "SPARQL dotazování" kdykoliv je detekována distribuce datové sady ve formě datové služby, která má jako specifikaci indikován protokol SPARQL (<https://www.w3.org/TR/sparql11-protocol/>) značící publikaci dat v 5* podobě.

SPARQL Endpoint	
Podmínky užití distribuce	Datová služba
Neobsahuje Autorské dílo ✓	Popis endpointu ✓ HTTP
Neobsahuje Originální databáze ✓	Endpoint SPARQL dotazování ✓ HTTP
Není chráněna Zvláštní právo pořizovatele databáze ✓	Specifikace ✓
Neobsahuje Osobní údaje 👤	

Asistent dotazování nad 5* daty

Asistent YASGUI je uživatelsky přívětivé rozhraní pro asistovanou tvorbu SPARQL dotazů a vizualizaci výsledků těchto dotazů. Rozhraní podporuje zvýrazňování syntaxe a automatické



doplňování prefixů IRI a IRI predikátů a tříd z dobře známých RDF slovníků. Umožňuje také dotazy sdílet pomocí zkrácených URI, např. <https://api.tripliedb.com/s/QdnSDtMJF>.

The screenshot shows a SPARQL query interface. The query is:

```
1 SELECT ?typ (COUNT(?s) AS ?pocet)
2 WHERE {
3   ?s a ?typ
4 }
5 GROUP BY ?typ
6 ORDER BY DESC(COUNT(?s))
7 LIMIT 10
```

The results table shows 10 entries:

typ	pocet
<https://slovník.gov.cz/agendový/104/pojem/působnost-v-agendě>	"764667"<http://www.w3.org/2001/XMLSchema#integer>
<https://slovník.gov.cz/legislativní/sbírka/111/2009/pojem/ustanovení-právního-předpisu>	"81565"<http://www.w3.org/2001/XMLSchema#integer>
<https://slovník.gov.cz/agendový/104/pojem/obslužný-kanál>	"80628"<http://www.w3.org/2001/XMLSchema#integer>
<https://slovník.gov.cz/agendový/104/pojem/úkon-služby-veřejné-správy>	"24021"<http://www.w3.org/2001/XMLSchema#integer>
<https://slovník.gov.cz/legislativní/sbírka/111/2009/pojem/pracoviště>	"20138"<http://www.w3.org/2001/XMLSchema#integer>
<https://slovník.gov.cz/agendový/104/pojem/oprávnění-k-přístupu-k-údaji>	"19874"<http://www.w3.org/2001/XMLSchema#integer>
<https://slovník.gov.cz/legislativní/sbírka/111/2009/pojem/osoba-stojící-v-čele-orgánu-veřejné-moci>	"19170"<http://www.w3.org/2001/XMLSchema#integer>
<https://slovník.gov.cz/legislativní/sbírka/111/2009/pojem/datová-schránka>	"18216"<http://www.w3.org/2001/XMLSchema#integer>
<https://slovník.gov.cz/legislativní/sbírka/111/2009/pojem/orgán-veřejné-moci>	"17993"<http://www.w3.org/2001/XMLSchema#integer>
<https://slovník.gov.cz/legislativní/sbírka/111/2009/pojem/role>	"17207"<http://www.w3.org/2001/XMLSchema#integer>

Na obrázku výše je vidět uživatelské rozhraní tak, jak vypadá po kliknutí na nový odkaz v NKOD u datové sady, která obsahuje distribuci ve formě datové služby typu SPARQL Endpoint s URL <https://rpp-opendata.egon.gov.cz/odrpp/sparql/>. Výchozí dotaz je:

```
SELECT ?typ (COUNT(?s) AS ?pocet)
WHERE {
  ?s a ?typ
}
GROUP BY ?typ
ORDER BY DESC(COUNT(?s))
LIMIT 10
```

Výchozí dotaz tedy zobrazuje 10 nejčastěji používaných tříd v daném SPARQL endpointu.

YASGUI je open-source software vyvíjený na GitHubu (<https://github.com/TriplyDB/Yasgui>) a volně licencovaný pomocí licence MIT. Lze tak využít buďto volně dostupnou instanci <https://yasgui.tripliedb.com/> a nebo nasadit vlastní kdekoli jinde.

Podoba katalogizačního záznamu datové sady publikované jako 5* data pomocí SPARQL endpointu

Jako příklad katalogizačního záznamu ve formátu JSON-LD, který umožní zobrazit odkaz na asistenta dotazování uvádíme zjednodušenou verzi [katalogizačního záznamu samotného](#)



[Národního katalogu otevřených dat](#) dle Otevřené formální normy Rozhraní katalogů otevřených dat: DCAT-AP-CZ, konkrétně část odpovídající distribuci v podobě SPARQL endpointu:

```
{
  "@context": "https://ofn.gov.cz/rozhraní-katalogů-otevřených-dat/2021-01-11/kontexty/rozhraní-katalogů-otevřených-dat.jsonld",
  "iri": "https://data.mvcr.gov.cz/zdroj/datové-sady/nkod/nkod-kompletní",
  "typ": "Datová sada",
  "distribuce": [
    {
      "typ": "Distribuce",
      "iri": "https://data.mvcr.gov.cz/zdroj/datové-sady/nkod/nkod-kompletní/distribuce/1",
      "název": {
        "cs": "Přístupový bod SPARQL",
        "en": "SPARQL endpoint"
      },
      "podmínky_užití": {
        "typ": "Specifikace podmínek užití",
        "autorské_dílo": "https://data.gov.cz/podmínky-užití/neobsahuje-autorská-díla/",
        "databáze_jako_autorské_dílo": "https://data.gov.cz/podmínky-užití/není-autorskopravně-chráněnou-databází/",
        "databáze_chráněná_zvláštními_právy": "https://data.gov.cz/podmínky-užití/není-chráněna-zvláštním-prámem-pořizovatele-databáze/",
        "osobní_údaje": "https://data.gov.cz/podmínky-užití/obsahuje-osobní-údaje/"
      },
      "přístupové_url": "https://data.gov.cz/sparql",
      "přístupová_služba": {
        "typ": "Datová služba",
        "přístupový_bod": "https://data.gov.cz/sparql",
        "popis_přístupového_bodu": "https://data.gov.cz/sparql",
        "iri": "https://data.mvcr.gov.cz/zdroj/datové-sady/nkod/nkod-kompletní/distribuce/1/přístupová-slужba",
        "název": {
          "cs": "Přístupový bod SPARQL",
          "en": "SPARQL endpoint"
        },
        "poskytuje_datovou_sadu": "https://data.mvcr.gov.cz/zdroj/datové-sady/nkod/nkod-kompletní",
        "specifikace": "https://www.w3.org/TR/sparql11-protocol/"
      }
    }
  ]
}
```

Klíčovou částí záznamu je indikace kompatibility datové služby s protokolem SPARQL, tedy "[specifikace](https://www.w3.org/TR/sparql11-protocol/)": "<https://www.w3.org/TR/sparql11-protocol/>" na objektu typu "Datová služba".

V registračním formuláři NKOD pak tato distribuce vypadá následovně:



Přístup k distribuci

Typ distribuce
↔ Datová služba

URL endpointu
🔗 <https://data.gov.cz/sparql>

Odkaz na popis endpointu. <https://data.gov.cz/sparql> [?](#) Odkaz na specifikaci služby (volitelné) <https://www.w3.org/TR/sparql11-protocol/> [?](#)

Název distribuce česky Přístupový bod SPARQL [?](#) Název distribuce anglicky (volitelné) SPARQL endpoint [?](#)

SMAŽ DISTRIBUCI

Testování použitelnosti

V této sekci se věnujeme testování použitelnosti prototypu pomocí standardní metodiky [System Usability Scale \(SUS\)](#). Testovací uživatelé nejprve prototyp vyzkoušeli na základě scénáře, a pak vyplnili dotazník založený na SUS. Pro plné využití potenciálu asistenta dotazování je třeba umět jazyk SPARQL.

Scénář

1. V uživatelském rozhraní Národního katalogu otevřených dat přistupte k detailu datové sady, která má distribuci ve formě SPARQL endpointu. Samotné hledání takové datové sady není součástí scénáře. Použít můžeme například datovou sadu [Agendy](#) od Ministerstva vnitra.
2. U distribuce datové sady ve formě SPARQL endpointu bude mít u sebe odkaz "SPARQL dotazování". Klikněte na něj.
3. Otevře se asistent pro dotazování jazykem SPARQL, který spustí předdefinovaný SPARQL dotaz.
4. Volitelně: Pokud SPARQL umíte, můžete dotaz libovolně změnit a využít tak automatické doplňování prefixů a našeptávání pojmů ze známých slovníků.

Výsledky dle metodiky SUS

Počet testovacích uživatelů: 3

Výsledné skóre a jeho interpretace: **63.33** - lehce podprůměrné skóre - jednomu z uživatelů prototyp evidentně nesedl.

Jednotlivé výsledky

Už./O.	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10	Skóre
1	3	4	4	3	4	4	4	4	4	1	87.5
2	0	0	0	1	0	2	2	0	0	2	17.5



3	4	4	4	2	4	4	3	4	2	3	85
---	---	---	---	---	---	---	---	---	---	---	----

Přílohy

Přílohou jsou zdrojové kódy a dokumentace uživatelského rozhraní NKOD (nástroj LinkedPipes DCAT-AP Viewer) s rozšířením o podporu asistenta dotazování ([dcat-ap-viewer-develop.zip](#)).