




jakubklimek Úprava před odevzdáním

 History

 1 contributor

 174 lines (147 sloc) | 16.1 KB

...

Uživatelská příručka - Národní katalog otevřených dat

1. ROZSAH PLATNOSTI A ÚČEL

Tato uživatelská příručka slouží k popisu uživatelské části Národního katalogu otevřených dat (NKOD), části projektu OD2.0.

2. DEFINÍCIA POJMOV A SKRATIEK

DCAT-AP-SK 2.0

Specifikace pro metadatový záznam datasetu otevřených dat

LKOD

Lokální katalog otevřených dat

NKOD

Národní katalog otevřených dat

POD

Portál otevřených dat

RDF

Resource Description Framework - datový model využívaný NKOD

SPARQL

Dotazovací jazyk nad daty v RDF

3. POPIS IS

NKOD slouží jako backend budoucího Portálu otevřených dat (POD). Uživatelské prostředí NKOD je tedy nyní minimalistické a tvoří ho jediná webová stránka.

Toto je testovacia verzia SPARQL Endpointu, ktorý bude súčasťou nového portálu otvorených dát. Zdrojom dát je stále súčasný data.gov.sk ktorý sa do NKODu harvestuje raz za deň.

opendata.mirri.tech

Slovenčina



SPARQL Endpoint pre Národný katalóg otvorených dát

Príklady dotazov

[100 datasetov a ich poskytovateľov](#)

100 datasetov a ich poskytovateľov

[Zoznam lokálnych katalógov údajov](#)

Zoznam lokálnych dátových katalógov a počty dátových sád v nich

[Počet datasetov na poskytovateľa](#)

Počet datasetov na poskytovateľa

[Dáta v CSV](#)

Datsety s distribúciami v CSV

[Nedostupnosť stiahnuteľných súborov](#)

Nedostupnosť stiahnuteľných súborov na základe meraní kvality

Zadaj SPARQL dotaz

Query

```
1 PREFIX dcterms: <http://purl.org/dc/terms/>
2 PREFIX dcat: <http://www.w3.org/ns/dcat#>
3 PREFIX foaf: <http://xmlns.com/foaf/0.1/>
4
5 SELECT ?mime_type (count (distinct ?distribuce) as ?počet_distribucí)
6 WHERE
7 {
8   ?datová_sada a dcat:Dataset ;
9   dcat:distribution ?distribuce .
10
11   ?distribuce dcat:mediaType ?mime_type.
12 }
13 GROUP BY ?mime_type
14 ORDER BY DESC(?počet_distribucí)
```

Table

Response

15 results in 0.05 seconds

Simple view

Ellipse

Filter query results

Page size: 50



mime_type	počet_distribucí
<http://www.iana.org/assignments/media-types/application/xml>	"4051"^^<http://www.w3.org/2001/XMLSchema#integer>
<http://www.iana.org/assignments/media-types/text/csv>	"3016"^^<http://www.w3.org/2001/XMLSchema#integer>
<http://www.iana.org/assignments/media-types/application/vnd.openxmlformats-officered+xml>	"1002"^^<http://www.w3.org/2001/XMLSchema#integer>
<http://www.iana.org/assignments/media-types/text/html>	"707"^^<http://www.w3.org/2001/XMLSchema#integer>
<http://www.iana.org/assignments/media-types/application/pdf;type=pdf1x>	"668"^^<http://www.w3.org/2001/XMLSchema#integer>
<http://www.iana.org/assignments/media-types/application/vnd.ms-excel>	"532"^^<http://www.w3.org/2001/XMLSchema#integer>
<http://www.iana.org/assignments/media-types/application/zip>	"341"^^<http://www.w3.org/2001/XMLSchema#integer>
<http://www.iana.org/assignments/media-types/application/geo+json>	"20"^^<http://www.w3.org/2001/XMLSchema#integer>
<http://www.iana.org/assignments/media-types/application/vnd.openxmlformats-officered+xml>	"18"^^<http://www.w3.org/2001/XMLSchema#integer>
<http://www.iana.org/assignments/media-types/application/xhtml+xml>	"16"^^<http://www.w3.org/2001/XMLSchema#integer>

```
10 <http://www.iana.org/assignments/media-types/application/json>
11 <http://www.iana.org/assignments/media-types/application/json>
12 <http://www.iana.org/assignments/media-types/image/tiff>
13 <http://www.iana.org/assignments/media-types/text/rtf>
14 <http://www.iana.org/assignments/media-types/text/plain>
15 <http://www.iana.org/assignments/media-types/application/msword>
```

Showing 1 to 15 of 15 entries

```
10 ...
"5"^^<http://www.w3.org/2001/XMLSchema...
"3"^^<http://www.w3.org/2001/XMLSchema...
"3"^^<http://www.w3.org/2001/XMLSchema...
"2"^^<http://www.w3.org/2001/XMLSchema...
"1"^^<http://www.w3.org/2001/XMLSchema...
```

< 1 >

[Našli ste na stránke chybu?](#)

Na webové stránke lze prepínat medzi 5 jazyky, vyberať predpripravené SPARQL dotazy a SPARQL dotazy spoušťať a vidieť výsledky.

Editor SPARQL dotazu je realizován nástrojem [Yasgui](#), jehož uživatelská dokumentace je [dostupná online](#).

Data v NKOD se řídí specifikací [DCAT-AP-SK 2.0](#). Data o výsledcích měření jsou pak založena na specifikaci [Data Quality Vocabulary](#).

Aktualizace NKOD probíhá zpravidla denně, v noci.

3.1 URL s výstupními daty

Kromě webové stránky samotné jsou vystavena i následující URL. Pro produkční prostředí je prefixem všech URL `https://opendata.mirri.tech`. Na toto prostředí existuje i redirect z `https://data.gov.sk/sparql`. Pro testovací prostředí je prefixem všech URL `https://opendata-test.mirri.tech`.

SPARQL endpoint NKOD

- [/api/sparql](#) - SPARQL endpoint NKOD

Soubory s obsahem NKOD v RDF

- [/subor/nkod.trig](#) - Dump obsahu NKOD v [RDF TriG](#)
- [/subor/lkody.trig](#) - Informace o registrovaných LKODEch v [RDF TriG](#)
- [/subor/nkod-metadata.ttl](#) - základní metadata o NKOD, zejména datum poslední aktualizace v [RDF Turtle](#)

Soubory s výsledky z měření kvality a dostupnosti

Dostupnost registrovaných zdrojů

- [/subor/kvalita/dostupnost/dostupnost.ttl](#) - Poslední detailní výsledky měření
- [/subor/kvalita/dostupnost/dostupnost-archiv.ttl](#) - Poslední agregované výsledky měření
- [/subor/kvalita/dostupnost-archiv/dostupnost-YYYY-MM-DD.ttl.gz](#) - Komprimovaný archiv agregovaných výsledků měření k danému datu

Dostupnost CORS na registrovaných zdrojích

- [/subor/kvalita/dostupnost-cors/cors.ttl](#) - Poslední detailní výsledky měření
- [/subor/kvalita/dostupnost-cors/cors-archiv.ttl](#) - Poslední agregované výsledky měření
- [/subor/kvalita/dostupnost-cors-archiv/dostupnost-cors-YYYY-MM-DD.ttl.gz](#) - Komprimovaný archiv agregovaných výsledků měření k danému datu

Kvalita metadatových záznamů

- [/subor/kvalita/kvalita/kvalita.ttl](#) - Poslední detailní výsledky měření
- [/subor/kvalita/kvalita/kvalita-archiv.ttl](#) - Poslední agregované výsledky měření
- [/subor/kvalita/kvalita-archiv/kvalita-YYYY-MM-DD.ttl.gz](#) - Komprimovaný archiv agregovaných výsledků měření k danému datu

Kombinované indikátory

- [/subor/kvalita/kombinace/kombinované-indikátory.ttl](#) - Poslední agregované indikátory
- [/subor/kvalita/kombinace-archiv/kombinované-indikátory-YYYY-MM-DD.ttl.gz](#) - Komprimovaný archiv k danému datu

Soubory s CSV reporty výsledků měření kvality a dostupnosti

Tyto soubory jsou v cestě `/subor/kvalita/indikatory/`.

- [a1-1-nedostupnost.csv](#) - Nedostupnost distribucí datových sad
- [a1-2-nedostupnost-seznam.csv](#) - Nedostupné distribuce datových sad
- [a1-3-nedostupnost-cors.csv](#) - Nedostupnost techniky CORS u distribucí ve formě souboru ke stažení
- [a1-4-nedostupnost-cors-seznam.csv](#) - Distribuce ve formě souboru ke stažení s nedostupnou technikou CORS u souboru ke stažení
- [a2-1-nedostupnost-schemat.csv](#) - Nedostupnost schémat distribucí datových sad
- [a2-2-nedostupnost-schemat-seznam.csv](#) - Nedostupná schémata distribucí datových sad
- [a2-3-nedostupnost-cors-schemat.csv](#) - Nedostupnost techniky CORS u schémat distribucí ve formě souboru ke stažení
- [a2-4-nedostupnost-schemat-cors-seznam.csv](#) - Distribuce ve formě souboru ke stažení s nedostupnou technikou CORS u schématu souboru ke stažení
- [a3-1-nedostupnost-podminek-uziti.csv](#) - Nedostupnost podmínek užití distribucí datových sad
- [a3-2-nedostupnost-podminek-uziti-seznam.csv](#) - Nedostupné podmínky užití distribucí datových sad
- [a4-1-nedostupnost-dokumentace.csv](#) - Nedostupnost uživatelské dokumentace datových sad
- [a4-2-nedostupnost-dokumentace-seznam.csv](#) - Nedostupné uživatelské dokumentace datových sad
- [a5-1-neshoda-media-type.csv](#) - Neshoda mezi formátem distribuce v NKOD a formátem indikovaným serverem - statistika
- [a5-2-neshoda-media-type-list.csv](#) - Neshoda mezi formátem distribuce v NKOD a formátem indikovaným serverem - seznam
- [a6-1-nedostupnost-endpoint-url.csv](#) - Nedostupnost přístupových bodů distribucí ve formě datové služby
- [a6-2-nedostupnost-endpoint-url-seznam.csv](#) - Nedostupné přístupové body distribucí ve formě datové služby
- [a6-3-nedostupnost-endpoint-url-cors.csv](#) - Nedostupnost techniky CORS u přístupových bodů distribucí ve formě datové služby
- [a6-4-nedostupnost-endpoint-url-cors-seznam.csv](#) - Přístupové body distribucí ve formě datové služby s nedostupnou technikou CORS
- [a7-1-nedostupnost-endpoint-description.csv](#) - Nedostupnost popisů přístupových bodů distribucí ve formě datové služby
- [a7-2-nedostupnost-endpoint-description-seznam.csv](#) - Nedostupné popisy

přístupových bodů distribucí ve formě datové služby

- [a7-3-nedostupnost-endpoint-description-cors.csv](#) - Nedostupnost techniky CORS u popisů přístupových bodů distribucí ve formě datové služby
- [a7-4-nedostupnost-endpoint-description-cors-seznam.csv](#) - Popisy přístupových bodů distribucí ve formě datové služby s nedostupnou technikou CORS
- [a8-1-nedostupnost-specifikaci-sluzeb.csv](#) - Nedostupnost specifikací datových služeb
- [a8-2-nedostupnost-specifikaci-sluzby-seznam.csv](#) - Nedostupné specifikace datových služeb
- [a9-1-nedostupnost-specifikace.csv](#) - Nedostupnost specifikací datových sad
- [a9-2-nedostupnost-specifikaci-seznam.csv](#) - Nedostupné specifikace datových sad
- [nkod-pocty-sad.csv](#) - Počty datových sad registrovaných z formuláře a přes LKOD
- [nkod-typy.csv](#) - Druhy registrace datových sad v NKOD
- [nkod-zebricek.csv](#) - Počty datových sad a distribucí dle poskytovatele
- [q1.csv](#) - Počet distribucí bez specifikovaných podmínek užití dle poskytovatele
- [q2.csv](#) - Počet datových sad, jejichž distribuce nemají specifikovány podmínky užití dle poskytovatele
- [q3.csv](#) - Počet záznamů datových sad nesplňujících povinné atributy dle poskytovatele
- [q3l.csv](#) - Záznamy datových sad nesplňujících povinné atributy dle poskytovatele
- [q4a.csv](#) - Media typy souborů ke stažení dle poskytovatele
- [q4b.csv](#) - Specifikace datových služeb dle poskytovatele
- [q4c.csv](#) - Formáty dat distribucí celkem
- [q5b.csv](#) - Počty poskytovatelů dle podmínek užití distribucí
- [q5c.csv](#) - Podmínky užití distribucí celkem
- [q5.csv](#) - Podmínky užití distribucí dle poskytovatele
- [q6c.csv](#) - Počty datových sad s danou periodicitou aktualizace celkem
- [q6.csv](#) - Počty datových sad s danou periodicitou aktualizace dle poskytovatele
- [q7a.csv](#) - Počty datových sad s daným klíčovým slovem
- [q7atop30.csv](#) - Počty datových sad s daným klíčovým slovem - prvních 30
- [q7b.csv](#) - Počty poskytovatelů datových sad s daným klíčovým slovem
- [q7btop30.csv](#) - Počty poskytovatelů datových sad s daným klíčovým slovem - prvních 30

- [q8.csv](#) - Počty distribucí datových sad s nevalidním MIME Typem
- [q81.csv](#) - Seznam distribucí datových sad s nevalidním MIME Typem

3.2 Registrační záznamy

V současnosti uživatelsky nelze spravovat registrační záznamy datových sad a lokálních katalogů. Tuto funkcionality přidá až Portál otevřených dat. Registrační záznamy tak spravuje Administrátor NKOD ručně v [GitHub repozitáři](#). Registrační záznam datové sady se řídí [DCAT-AP-SK 2.0](#). Registrační záznam LKOD typu [DCAT-AP Dokumenty](#) vypadá například takto:

```
@prefix dct: <http://purl.org/dc/terms/> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .

@prefix lkod: <https://data.gov.sk/def/local-catalog-type/> .

<https://data.gov.sk/set/catalog/0144444444> a dcat:Catalog, lkod:1 ;
  dct:title "DCAT-AP-SK 2.0 DCAT-AP Dokumenty LKOD 1"@sk, "DCAT-AP-SK 2.0 DCAT-
  dcat:contactPoint [
    a vcard:Organization ;
    vcard:fn "Jakub Klímek"@sk ;
    vcard:hasEmail <mailto:jakub.klimek@matfyz.cuni.cz>
  ];
  dct:publisher <https://data.gov.sk/id/legal-subject/00681156> ;
  dcat:endpointURL <https://raw.githubusercontent.com/jakubklimek/sk-lkod-1/main> ;
  foaf:homepage <https://github.com/jakubklimek/sk-lkod-1> .
```

kde `lkod:1` indikuje typ LKOD dle číselníku typů LKOD.

Pro přidání registračního záznamu doporučujeme zkopírovat nějaký existující a upravit IRI a obsažená metadata.

4. NÁROKY NA POUŽÍVATEĽA

Uživatelé NKOD jsou 3 druhů.

Návštěvník bez znalosti SPARQL

Tento si vystačí s předpřipravenými SPARQL dotazy a tedy nemusí mít žádné zvláštní dovednosti.

Návštěvník se znalostí SPARQL

Pokud chce návštěvník psát vlastní SPARQL dotaz, musí ovládat SPARQL a RDF, znát specifikaci DCAT-AP-SK 2.0, a pro zpracovávání údajů z měření kvality pak ještě Data Quality Vocabulary.

Externí IS

Externí IS bude přistupovat k webové službě SPARQL endpointu. Potřebuje knihovnu pro práci se SPARQLelem, znát RDF, znát specifikaci DCAT-AP-SK 2.0, a pro zpracovávání údajů z měření kvality pak ještě Data Quality Vocabulary.

[Give feedback](#)