



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost



MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY

C3V3

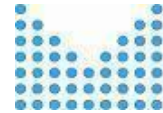
Zajištění podpory publikace výsledků Výroční zprávy ke stavu otevřených dat ve VS ČR v POD

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

Verze výstupu: 01



Zobrazení metrik kvality relevantních pro výroční zprávu a jejich vizualizace

V této sekci jsou ukázány vizualizace jednotlivých metrik z C2V9 a obecných statistik NKOD relevantních pro výroční zprávu.

Počet poskytovatelů a datových sad v NKOD

Tato vizualizace ukazuje počet poskytovatelů v NKOD, počet datových sad za jednotlivé poskytovatele a počet datových sad v NKOD celkem.



Soulad záznamů datových sad s Otevřenou formální normou Rozhraní katalogů otevřených dat DCAT-AP-CZ

Tato vizualizace ukazuje, kolik poskytovatelů poskytuje záznamy datových sad do Národního katalogu otevřených dat v souladu s Otevřenou formální normou (OFN) [Rozhraní katalogů otevřených dat: DCAT-AP-CZ](#) vyvinutou v rámci C2V1. Dále pak zobrazuje počet datových sad jednotlivých poskytovatelů, jejichž popis OFN nespňuje a počet takových datových sad celkem.

V druhé části je pak zobrazeno, kolik poskytovatelů používá již nepodporované rozhraní Lokálních katalogů otevřených dat CKAN API.





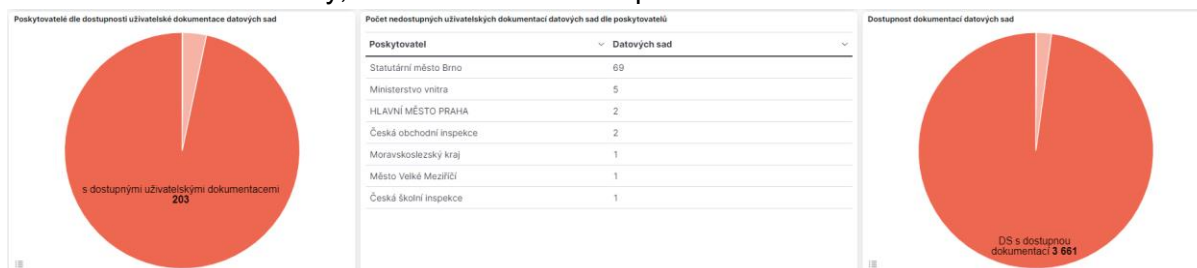
Podmínky užití datových sad - specifikace a dostupnost

Tato vizualizace ukazuje, kolik poskytovatelů má registrovány datové sady bez podmínek užití a kolik je takových datových sad v NKOD celkem. Dále ukazuje, kolik poskytovatelů a datových sad z těch, které specifikovány podmínky užití má, tak je ale nemá dostupné.



Dokumentace datových sad - specifikace a dostupnost

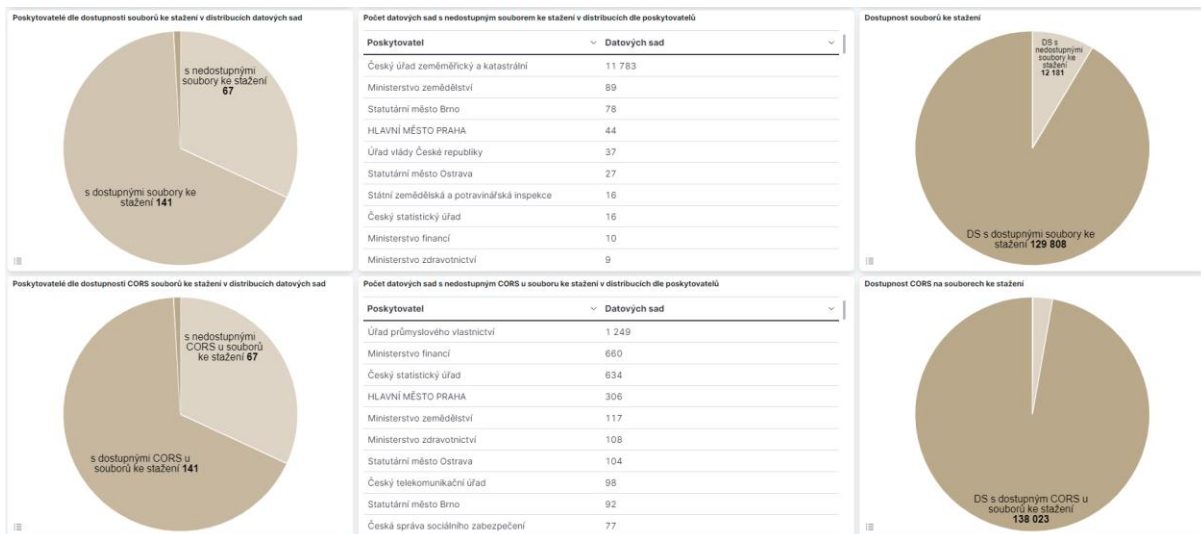
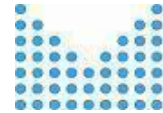
Tato vizualizace ukazuje, kolik poskytovatelů a jejich datových sad má registrovanou dokumentaci datové sady, která však není dostupná.



Distribuce ve formě souborů ke stažení - dostupnost a dostupnost techniky CORS

Tato vizualizace ukazuje, kolik poskytovatelů a jejich datových sad má distribuce ve formě souboru ke stažení, který však

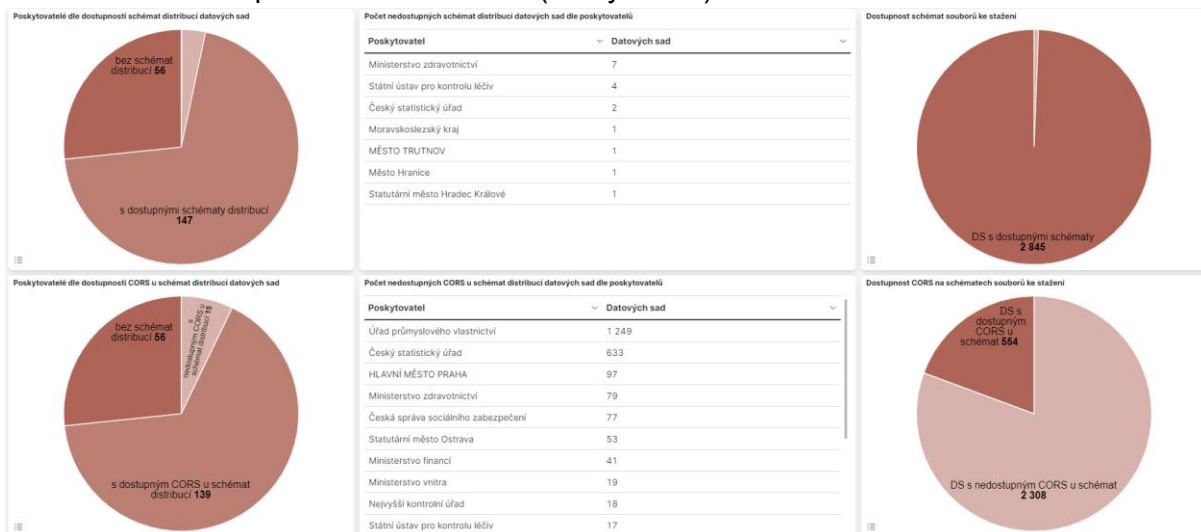
1. není dostupný (první řádek)
2. nemá dostupnou techniku CORS (druhý řádek).



Schémat souborů ke stažení - dostupnost a dostupnost techniky CORS

Tato vizualizace ukazuje, kolik poskytovatelů a jejich datových sad má schéma souboru ke stažení, které však

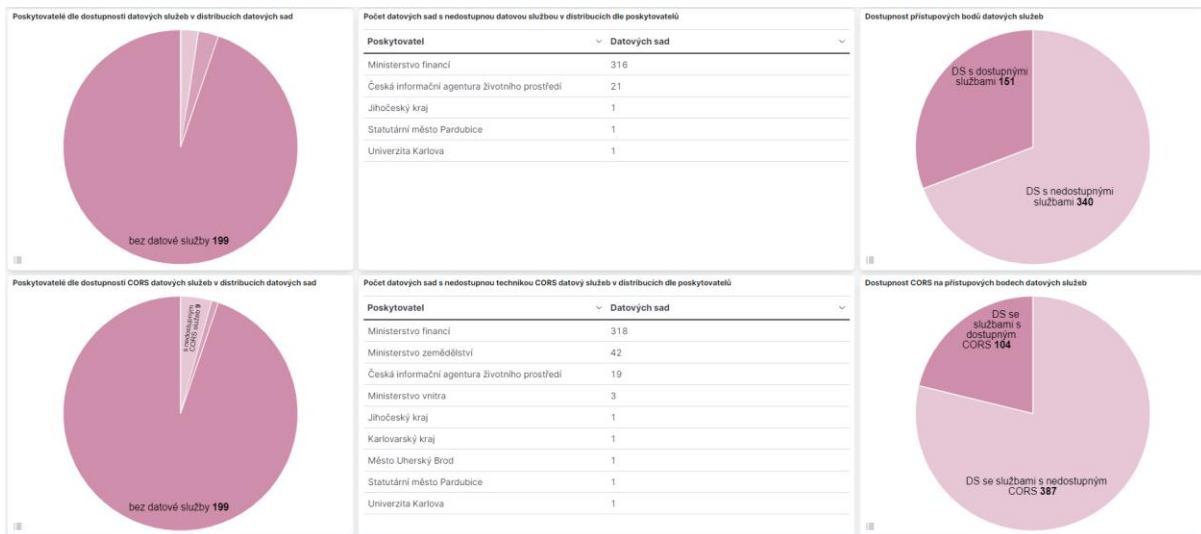
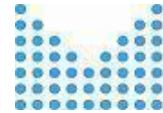
1. není dostupné (první řádek)
2. nemá dostupnou techniku CORS (druhý řádek).



Přístupové body datových služeb - dostupnost a dostupnost techniky CORS

Tato vizualizace ukazuje, kolik poskytovatelů a jejich datových sad má distribuci ve formě datové služby, jejíž přístupový bod však

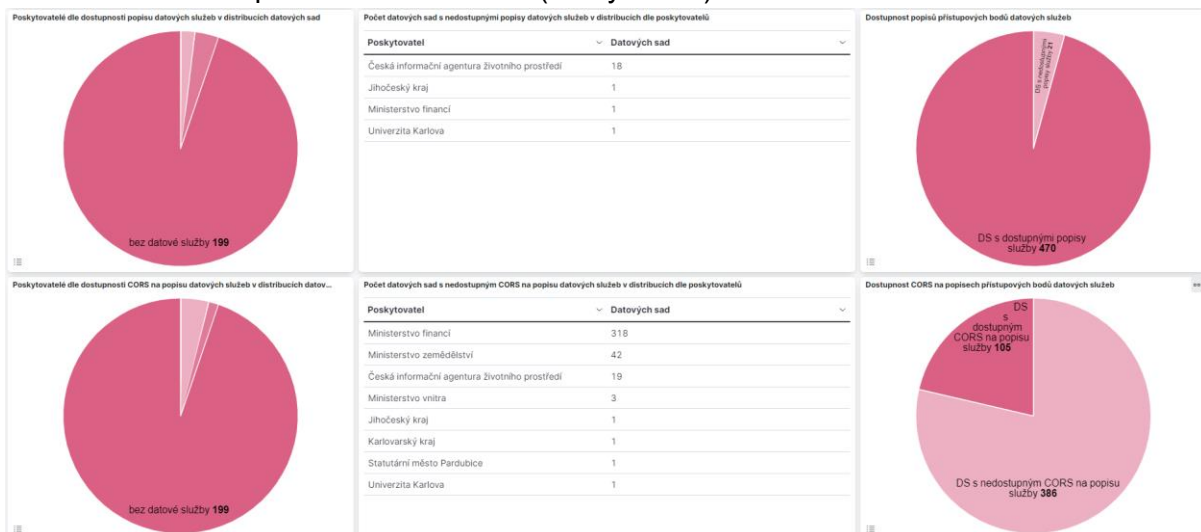
1. není dostupný (první řádek)
2. nemá dostupnou techniku CORS (druhý řádek).



Popisy datových služeb - dostupnost a dostupnost techniky CORS

Tato vizualizace ukazuje, kolik poskytovatelů a jejich datových sad má distribuci ve formě datové služby, jejíž popis přístupového bodu však

1. není dostupný (první řádek)
2. nemá dostupnou techniku CORS (druhý řádek).



Počty datových sad indikujících soulad s otevřenými formálními normami

Tato vizualizace ukazuje, kolik datových sad indikuje, že jsou v souladu s otevřenými formálními normami vyvíjenými v rámci výstupu C2V1.



Počet datových sad dle otevřených formálních norem

Export

OFN	Počet datasetů
https://ofn.gov.cz/prostorová-data/2019-08-22/	137 262
https://ofn.gov.cz/úřední-desky/2021-07-20/	397
https://ofn.gov.cz/úřední-desky/	360
https://ofn.gov.cz/číselníky/draft/	42
https://ofn.gov.cz/číselníky/	22
https://ofn.gov.cz/číselníky/2022-02-08/	22
https://ofn.gov.cz/%C3%BA%CE%99edn%C3%AD-desky/2021-07-20/	3
https://ofn.gov.cz/úřední-desky	2
https://ofn.gov.cz/%C3%BA%CE%99edn%C3%AD-desky	1
https://ofn.gov.cz/%C3%BA%CE%99edn%C3%AD-desky/	1
https://ofn.gov.cz/propojená-data/	1

Technické řešení

V této sekci se věnujeme technickému řešení tvorby popsaných interaktivních vizualizací.

Požadavky

Požadavky na technické řešení byly následující:

1. Možnost generovat vizualizace na základě otevřených dat záznamů z NKOD
2. Open-source software s grafickým UI
3. Možnost nasazení na servery MV ČR běžících v prostředí UPAAS

Zvolené řešení

Jako základ řešení byl zvolen [ELK stack](#), tj. index [Elasticsearch](#) a vizualizační systém [Kibana](#). Dále byl implementován software pravidelně doplňující Elasticsearch index na základě indikátorů kvality měřených v rámci výstupu C2V9. Výsledkem je software [Dashboard Indexer](#) publikovaný na GitHubu jako open source. GitHub repozitář obsahuje instrukce pro instalaci pomocí systému [Docker](#).

Systémové požadavky

Zvolené řešení vyžaduje (virtuální) stroj se 4 jádry, alespoň 8 GB paměti RAM, 20 GB místa na disku, instalovaným OS Linux (preferované je Ubuntu) a systémem Docker a docker-compose. Řešení je testováno na verzi Ubuntu 21.10.

Konfigurace

Pro konfiguraci vizualizéru Kibana je třeba použít příložený JSON soubor `export.ndjson`. Pro konfiguraci Dashboard Indexeru je třeba použít příložený JSON soubor `indexer-configs.conf`.

Integrace do Portálu otevřených dat

Bylo přidáno tlačítko “Dashboard NKOD” do [sekce Datová kvalita](#) v Portálu otevřených dat ukazující na URL dashboardu “Aktuální NKOD” v Kibana.



Dashboard zobrazuje aktuální situaci k poslednímu dni měření kvality a dostupnosti jednotlivých datových sad a indikátorů agregovaných na úroveň jednotlivých poskytovatelů.

OTEVŘENÁ DATA




Vzdělávání Články Datová kvalita Národní katalog otevřených dat Další informace

Datová kvalita (nejen) v oblasti otevřených dat

Datová kvalita je obecně velmi široký pojem. Vystihnout se však dá jako "přípravenost dat k použití". Nízká datová kvalita tedy představuje zásadní problém pro využívání potenciálu otevřených dat.

Za kvalitu svých dat i svých katalogizačních záznamů zodpovídají jednotliví poskytovatelé. Ti tedy musí na kvalitu svých dat dbát, pravidelně ji sledovat a dělat kroky k nápravě zjištěných nedostatků. Zde nabízíme přehled nástrojů, kterými poskytovatele ve zlepšování kvality jejich dat podporujeme.

Nástroje pro zlepšení kvality otevřených dat

Statistiky NKOD	Příklady špatné praxe	Otevřené formální normy
 <p>Denně měříme dostupnost zdrojů registrovaných v Národním katalogu otevřených dat spolu se statistikami kvality jejich metadatových záznamů. Mají právě vaše datové sady vyplněné všechny povinné položky? Jsou vámi registrované soubory a služby dostupné?</p> <p>Statistiky NKOD Dashboard NKOD</p>	 <p>Sbíráme příklady špatné praxe v otevírání dat a postupy jejich nápravy. Všem tak pomáháme se učit z chyb ostatních. Také ve formátu CSV oddělujete hodnoty středníkem? Dělíte distribuce datové sady dle časového pokrytí nebo dle tématu?</p> <p>Špatná praxe</p>	 <p>Tvoříme otevřené formální normy - standardy, jejichž dodržování přispívá ke zvyšování interoperability otevřených dat. Otevřené formální normy se týkají jak obecných oblastí, jako propojená data či geodata, tak i konkrétních datových sad jako turistické cíle či aktuality.</p> <p>Otevřené formální normy</p>

Pro účely tvorby výroční zprávy lze navíc nastavit požadované datum, ke kterému mají být informace zobrazeny. To může být třeba datum, ke kterému je výroční zpráva vydávána. Tyto vizualizace lze pak přímo použít jako součást výroční zprávy, a lze získat URL dashboardu ke konkrétnímu datu, které pak lze použít jako interaktivní vizualizaci dat z výroční zprávy.