

Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát 2020

na základe

Akčného plánu Iniciatívy pre otvorené vládnutie v Slovenskej republike
na roky 2017 – 2019, schváleného Uznesením vlády SR č. 104/2017,
Akčného plánu Iniciatívy pre otvorené vládnutie v Slovenskej republike
na roky 2020 – 2021, schváleného Uznesením vlády SR č. 553/2019

Bratislava, 29.06.2020



Pre Úrad splnomocnenca vlády SR pre rozvoj občianskej spoločnosti spracoval:
Martin Šechný, Spoločnosť pre otvorené informačné technológie

Obsah

| | |
|--|----|
| Úvod | 3 |
| 1 Kvantitatívna analýza | 4 |
| 1.1 Počet zverejnených datasetov | 4 |
| 1.2 Tagy (kľúčové slová) | 6 |
| 1.3 Licencie datasetov | 7 |
| 1.4 Formáty datasetov | 8 |
| 1.5 Aktualizácia datasetu | 9 |
| 1.6 Veľkosť datasetu | 9 |
| 1.7 Metadáta k datasetu | 9 |
| 1.8 Používatelia datasetov | 10 |
| 2 Obsah zverejňovaných datasetov | 11 |
| 2.1 Publikačné minimum | 11 |
| 2.2 Prieskum dopytu verejnosti po najžiadanejších datasetoch a API | 12 |
| 2.3 Správa o implementácii akčného plánu 2017 – 2019 | 17 |
| 2.4 Užitočnosť otvorených dát pre používateľa | 18 |
| 3 Kvalitatívna analýza | 19 |
| 3.1 Hodnotenie datasetov | 19 |
| 3.2 Global open data index | 21 |
| 3.3 Open data barometer | 21 |
| 3.4 Open data watch | 21 |
| 3.5 Index digitálnej ekonomiky a spoločnosti (DESI) | 22 |
| 3.6 Open data maturity index | 22 |
| 3.7 Nástroje | 22 |
| 3.8 Aplikácie | 23 |
| 3.9 API | 23 |
| Záver | 24 |

Úvod

Slovenská republika sa v roku 2011 zapojila do medzinárodnej **Iniciatívy pre otvorené vládnutie**. Úrad splnomocnenca vlády SR pre rozvoj občianskej spoločnosti zastrešuje túto iniciatívu krátkodobými akčnými plánmi. V tomto období aktuálny **Akčný plán Iniciatívy pre otvorené vládnutie v Slovenskej republike na roky 2020 – 2021**, schválený Uznesením vlády SR č. 553/2019, v kapitole 1 *Otvorené informácie* obsahuje úlohu, z ktorej vyplýva potreba každoročnej analýzy zverejňovania datasetov ústredných orgánov štátnej správy:

1.2 *Pravidelne zverejňovať datasety a aplikačné programové rozhrania (API) v gescii ministerstiev a nimi zriadených organizácií a ostatných ústredných orgánov štátnej správy.*

Vypracovať analýzu (každoročne) je zároveň pravidelnou úlohou B.10 z *Akčného plánu Iniciatívy pre otvorené vládnutie v Slovenskej republike na roky 2017 – 2019*, schváleného Uznesením vlády SR č. 104/2017. Cieľom zverejňovania otvorených dát verejnej správy je zvyšovanie transparentnosti a podpora hospodárstva založeného na dátach. Ústredné orgány štátnej správy zhromažďujú a spravujú veľké množstvo dát, ktoré je možné sprístupniť (okrem dát, kde existujú zákonné prekážky), a to s verejnou licenciou a v strojovo spracovateľných formátoch. Ku každému zverejnenému datasetu majú byť pripojené metadáta, ktoré opisujú obsah a štruktúru datasetu. Na zverejňovanie otvorených dát sa vzťahuje tiež Národná koncepcia informatizácie verejnej správy¹, Stratégia a akčný plán sprístupnenia a používania otvorených údajov verejnej správy².

Predchádzajúce analýzy zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát boli realizované za roky 2017, 2018, 2019. Táto analýza mapuje zverejnené datasety ústredných orgánov štátnej správy na portáli otvorených dát (data.gov.sk), posudzuje zmapované datasety z kvantitatívneho aj kvalitatívneho hľadiska, porovnáva vývoj zverejňovania datasetov v čase. Táto analýza nemá skúmať zverejnené datasety verejnej správy mimo ústredných orgánov štátnej správy, ani zverejnené datasety územnej samosprávy, pre tie sa uvádzajú orientačné počty.



Obrázok 1: <https://data.gov.sk>

Správcom portálu data.gov.sk je *Národná agentúra pre sieťové a elektronické služby* (NASES), ktorá patrí pod *Úrad podpredsedu vlády SR pre investície a informatizáciu* (ÚPVII), ktorý je v postavení ústredného orgánu štátnej správy tvorcom politiky v oblasti informatizácie verejnej správy a správy informačných technológií verejnej správy. Portál poskytuje tiež vyhľadávanie a filtrovanie datasetov podľa rôznych kritérií, návody a iné súvisiace informácie.

1 <https://www.vicepremier.gov.sk/sekcie/narodna-koncepcia-informatizacie-verejnej-spravy-nikvs/index.html>

2 https://www.slovensko.sk/img/CMS4/strategia_otvorene_udaje.pdf

1 Kvantitatívna analýza

V tejto kapitole **sumarizujeme počty datasetov podľa viacerých kritérií** a porovnávame podľa rokov zverejnenia. Najprv uvádzame tabuľku s absolútnymi počtami datasetov podľa organizácií, potom grafy celkového počtu datasetov, počtu datasetov podľa tagov, počtu datasetov podľa typu licencie, počtu datasetov podľa formátov, ďalej graf porovnania aktuálnosti datasetov a ešte počty datasetov podľa veľkosti.

1.1 Počet zverejnených datasetov

Je možné zistiť rozdielne **absolútne počty datasetov** pre dané prípady **v závislosti od vyhl'adavacieho kritéria a zvolenej metodiky**. Na portáli otvorených dát sa nachádzajú viaceré datasety s rovnakým názvom od rovnakej organizácie, líšia sa v zdrojových dátach (napr. obsahujú rôzne kategórie dát alebo obsahujú dáta z rôznych období). Niektoré datasety sú zverejnené vo viacerých formátoch. Dataset s daným názvom a od danej organizácie sa považuje aktuálne za jeden, bez ohľadu na zdrojové dáta, obdobia, formát datasetu. Napriek tomu napr. dataset *Vestník verejného obstarávania*³ od Úradu pre verejné obstarávanie je vždy nanovo zverejňovaný mesačne, aj keď ide o rovnaký dataset, a to po častiach za rôzne (nerovnaké) obdobia, spravidla denne. Aj metodika analýz z predchádzajúcich rokov sa môže líšiť, preto absolútne počty datasetov nemožno považovať za presné, sú skôr orientačné. Dáta v tabuľke 1 boli zisťované k dátumom⁴: 27.08.2013, 08.03.2015, 31.03.2017, 30.06.2018, 25.06.2019, 20.05.2020. Portál data.gov.sk v záložke Štatistiky uvádza počet organizácií, sumárny počet datasetov⁵.

Zoznam ústredných orgánov štátnej správy je prevzatý z portálu www.slovensko.sk⁶, sú uvedené v tabuľke na bielych riadkoch. Táto analýza nemá skúmať datasety ostatných organizácií štátnej a verejnej správy mimo ústredných orgánov štátnej správy, avšak na portáli data.gov.sk sa vyskytujú datasety aj od týchto organizácií, preto ich uvádzame v tabuľke na sivých riadkoch. Táto analýza nemá skúmať ani zverejnené datasety územnej samosprávy, avšak niektoré sú zverejňované na portáli data.gov.sk, preto ich uvádzame sumárne v tabuľke na modrom riadku. Na portáli data.gov.sk sa vyskytujú datasety od niekoľkých organizácií patriacich pod príslušné ústredné orgány štátnej správy samostatne, tie sme prirátali k príslušným ústredným orgánom štátnej správy. Zoznam priradených organizácií je na portáli www.slovensko.sk⁷, ale tento zoznam zrejme nerešpektuje zriaďovateľskú pôsobnosť nadradených organizácií ku niektorým podriadeným organizáciám. Nesprávne priradenie organizácie môže byť zdrojom čiastkovej chyby v absolútnych počtoch datasetov, najmä v minulých rokoch. Súčet všetkých datasetov v roku sa týmto nezmení.

Nasledujúcu stranu vyplní tabuľka absolútnych počtov zverejnených datasetov sumárne za ústredné orgány štátnej správy s ich podriadenými organizáciami (aj sumárne za vybrané ostatné orgány štátnej a verejnej správy a územnej samosprávy), v porovnaní podľa rokov, za ktoré bola robená analýza. Dáta z roku 2020 sú k 20.05.2020. Portál data.gov.sk v záložke Štatistiky uvádza počet organizácií 86 a celkový počet datasetov 2 304 (líši sa od súčtu v tabuľke).

3 Napr. <https://data.gov.sk/dataset/vestnik-verejneho-obstaravania-20205>

4 <https://data.gov.sk/data.gov.sk/organization>

5 Sumárne počty v štatistike data.gov.sk (na obrázku) sa nezodujú so skutočným počtom datasetov po kliknutí.

6 <https://www.slovensko.sk/sk/agendy/agenda/ministerstva-a-ostatne-ustredne1>

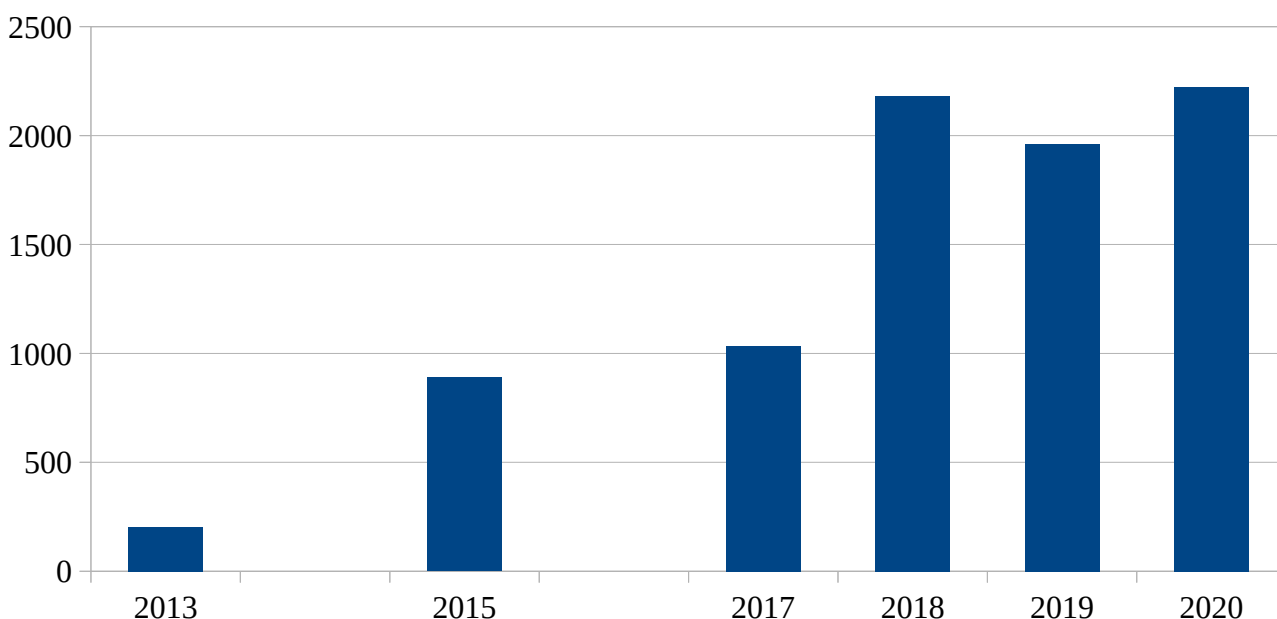
7 <https://www.slovensko.sk/sk/institucie>

Tabuľka 1: Počet zverejnených datasetov podľa organizácií a rokov

| Organizácia | 2013 | 2015 | 2017 | 2018 | 2019 | 2020 |
|--|------------|------------|--------------|--------------|--------------|--------------|
| Ministerstvo dopravy a výstavby SR | 8 | 10 | 10 | 32 | 66 | 60 |
| Ministerstvo financií SR | 11 | 22 | 22 | 20 | 13 | 14 |
| Ministerstvo hospodárstva SR | 7 | 8 | 9 | 17 | 17 | 34 |
| Ministerstvo kultúry SR | 10 | 11 | 11 | 35 | 25 | 26 |
| Ministerstvo obrany SR | 3 | 4 | 4 | 19 | 8 | 8 |
| Ministerstvo pôdohospodárstva a rozvoja vidieka SR | 65 | 68 | 67 | 209 | 457 | 469 |
| Ministerstvo práce, sociálnych vecí a rodiny SR | 5 | 5 | 5 | 5 | 13 | 13 |
| Ministerstvo spravodlivosti SR | 5 | 8 | 8 | 17 | 16 | 17 |
| Ministerstvo školstva, vedy, výskumu a športu SR | 5 | 24 | 24 | 34 | 26 | 26 |
| Ministerstvo vnútra SR | 24 | 40 | 41 | 253 | 61 | 62 |
| Ministerstvo zahraničných vecí a európskych záležitostí SR | 14 | 14 | 16 | 25 | 25 | 11 |
| Ministerstvo zdravotníctva SR | 9 | 18 | 95 | 184 | 262 | 381 |
| Ministerstvo životného prostredia SR | 9 | 9 | 22 | 41 | 25 | 42 |
| Národný bezpečnostný úrad | | | | 4 | 23 | 23 |
| Protimonopolný úrad SR | | | | 7 | 7 | 10 |
| Správa štátnych hmotných rezerv SR | | | | 5 | 7 | 10 |
| Štatistický úrad SR | 9 | 586 | 606 | 1 101 | 665 | 745 |
| Úrad geodézie, kartografie a katastra SR | 2 | 2 | 2 | 2 | 2 | 5 |
| Úrad jadrového dozoru SR | | | | 6 | 7 | 7 |
| Úrad podpredsedu vlády SR pre investície a informatizáciu | 0 | 23 | 12 | 36 | 39 | 36 |
| Úrad pre normalizáciu, metrologiu a skúšobníctvo SR | | | | 3 | 7 | 7 |
| Úrad pre verejné obstarávanie | 7 | 7 | 46 | 16 | 79 | 90 |
| Úrad priemyselného vlastníctva SR | | | | 19 | 8 | 9 |
| Úrad vlády SR | 7 | 11 | 11 | 13 | 15 | 18 |
| Generálna prokuratúra SR | | | | 8 | 8 | 8 |
| Kancelária prezidenta SR | | | | | | |
| Kancelária súdnej rady SR | | | | | | |
| Kancelária verejného ochrancu práv | | | | | | |
| Národná rada SR | | | | | 0 | 10 |
| Najvyšší kontrolný úrad SR | | | | | 2 | 10 |
| Národná banka Slovenska | | | | | | |
| Sociálna poisťovňa | | | | | | |
| Úrad na ochranu osobných údajov SR | | | | | | |
| Úrad pre reguláciu sieťových odvetví | | | | | | |
| Úrad pre Slovákov žijúcich v zahraničí | | | | | | |
| Ústav pamäti národa | | | | | | |
| Ústavný súd SR | | | | | | |
| (nezaradená) | 4 | 1 | 1 | 3 | 11 | |
| organizácie územnej samosprávy spolu | 0 | 18 | 22 | 67 | 68 | 71 |
| Spolu | 204 | 889 | 1 034 | 2 181 | 1 962 | 2 222 |

V absolútnom počte zverejnených datasetov ústredných orgánov štátnej správy vedie Štatistický úrad SR, čo je pochopiteľné. Za zamyslenie stojí fakt, že viaceré ostatné orgány štátnej a verejnej správy nemajú zverejnený ani jeden dataset na portáli data.gov.sk (neskúmame zverejňovanie informácií na iných webových sídlach). Z územnej samosprávy dlhodobo vedie mesto Prešov.

Absolútne počty zverejnených datasetov by v princípe mali s časom rásť. Môže sa vyskytnúť pokles, ak organizácia vymaže zverejnený dataset (napr. kvôli chybným dátam alebo chybným metadáтам alebo nevhodnej licencií), alebo ak viaceré zverejnené datasety sú neskôr zjednotené do jedného (napr. jeden dataset s viacerými zdrojovými dátami alebo za viaceré obdobia).



Graf 1: Celkový počet datasetov podľa rokov

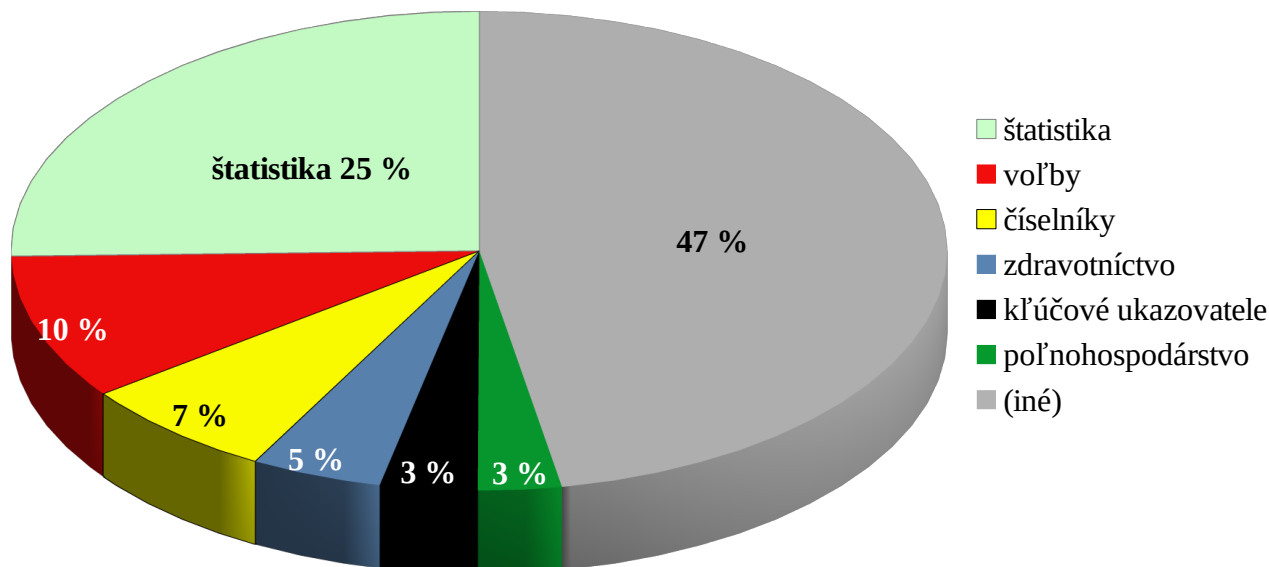
Poznámka: Celkový počet datasetov v jednotlivých rokoch je skôr orientačný.

1.2 Tagy (kľúčové slová)

Portál data.gov.sk v záložke Datasety umožňuje filtrovať datsety podľa tagu (kľúčového slova) pre ľahšie vyhľadávanie⁸, čo je veľmi užitočná funkcia. Portál ponúka niekoľko obľúbených, často používaných tagov. Tieto počty v prepočte na percentá uvádzame vo forme grafu. Celkový počet datasetov zistený podľa tagu sa líši od celkového počtu datasetov, zisteného podľa organizácie, pretože daný dataset môže mať priradených aj viac tagov.

Najväčší podiel majú datasety s tagom *štatistika*, čo je pochopiteľné, keďže Štatistický úrad SR má najväčší počet zverejnených datasetov. Na druhom mieste sú datasety s tagom *voľby*. Záujem používateľov o datasety volieb zrejme bude časovo variabilný podľa dátumu volieb. Na treťom mieste sú datasety s tagom *číselníky*. Tieto datasety sú dôležité pre vývojárov softvéru, lebo číselník predstavuje referenčný zoznam istého typu dát použiteľný cez API (napr. zoznam krajín, zoznam obcí, terminologický zoznam pre danú oblasť).

⁸ https://data.gov.sk/dataset? tags_limit=0

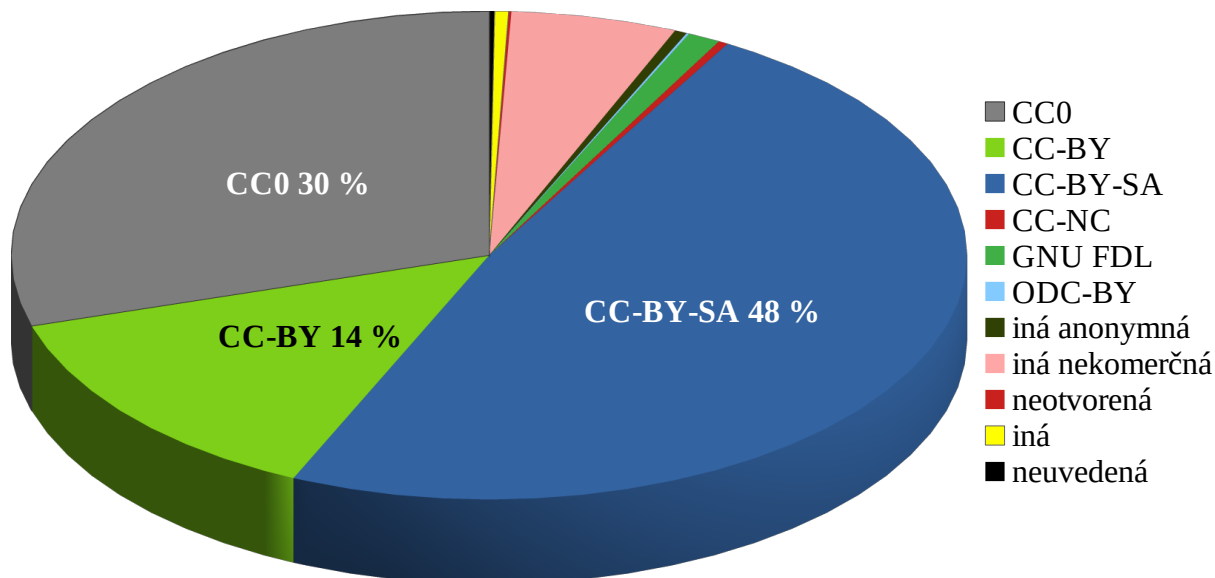


Graf 2: Počet datasetov podľa tagov (kľúčových slov)

Tagy nie sú definované jednotne, sú vo viacerých tvaroch, sú písané malými aj veľkými písmenami.

1.3 Licencie datasetov

Portál data.gov.sk v záložke Datasets umožňuje filtrovať datasety podľa typu licencie. Uvedené počty datasetov po prepočte na percentá uvádzame vo forme grafu. Celkový počet datasetov pri licenciách sa nezhoduje s celkovým počtom datasetov zisteným podľa organizácií.



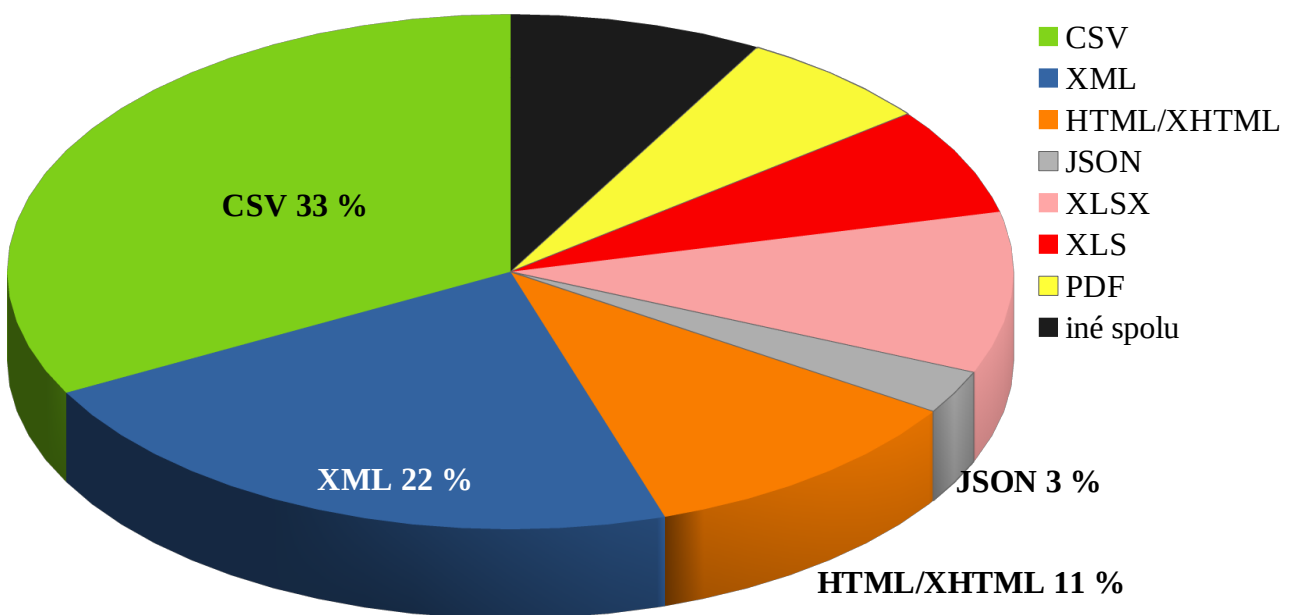
Graf 3: Typy licencií

Poradie najpoužívanejších licencií CC-BY-SA, CC0, CC-BY sa oproti roku 2019 nezmenilo. Napriek tomu, že portál data.gov.sk plní funkciu portálu otvorených dát, sú zverejnené aj datasety bez uvedenia licencie alebo s obmedzujúcou licenciou, čo je v rozpore so štandardami⁹. Odporúčané licencie sú CC0, CC-BY. Poskytovateľ datasetov, ktorý neuvádza licenciu, je Úrad vlády SR.

⁹ Vyhláška ÚPVII č. 78/2020 Z. z. (Vyhláška o štandardoch pre informačné technológie verejnej správy), § 39, <https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2020/78/20200501>

1.4 Formáty datasetov

Portál data.gov.sk v záložke Datasets umožňuje filtrovať datasety podľa formátu súboru. Uvedené počty datasetov v prepočte na percentá uvádzame v grafe. Celkový počet datasetov zistený pri formátoch je orientačný, nesedí s celkovým počtom datasetov zisteným pri organizáciách, pretože niektoré datasety sú zverejnené vo viacerých formátoch a formáty nie sú jednoznačne definované. Niektoré formáty uvedené na portáli nie sú formátmi skutočných datasetov, ale pomocnými súbormi k datasetom (napr. XSD ku XML). Niektoré formáty sú nesprávne alebo duplicitne pomenované. Medzi formáty sú zaradené aj hypertextové odkazy na iné sídlo mimo data.gov.sk.



Graf 4: Formáty datasetov

V roku 2018 dominoval formát XML. V rokoch 2019 a 2020 je najpoužívanejším formát CSV, nasledovaný formátmi XML, HTML/XHTML, JSON. Formáty CSV, XML, JSON sú najvhodnejšie pre zverejňovanie strojovo spracovateľných tabuľkových dát s textovými a číselnými dátami. Formáty HTML, XHTML sú vhodné na zverejňovanie rôznorodých dát, ľahko prístupných pre človeka, ale sú ťažšie spracovateľné strojom, nie sú vhodné pre otvorené dáta.¹⁰

Stále veľa datasetov je zverejňovaných i v ďalších formátoch, ktoré majú horšiu strojovú spracovateľnosť, najčastejšie XLSX, XLS, PDF. Relatívne malý počet datasetov je sprístupnený v niektorom zastaranom alebo inom nevhodnom formáte, napr. XLS, RTF, DOCX, DOC, RAR.

Filter podľa formátov na portáli data.gov.sk v záložke Datasets by mal obsahovať každý formát samostatne a jednoznačne. Filter zobrazuje tiež duplicitné alebo nesprávne definované formáty, napr. „xlsx“, „xml, xsl“, „api, csv“.

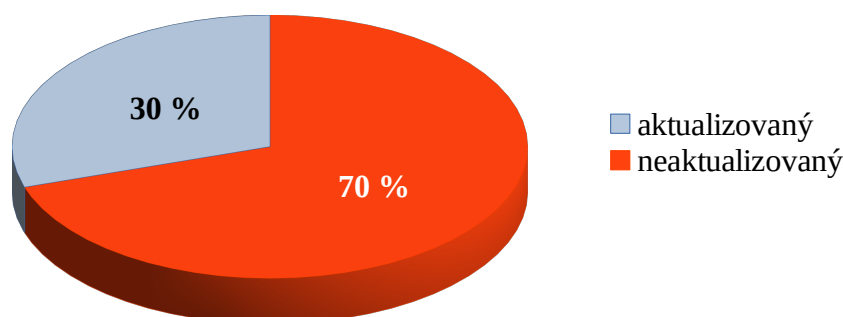
Datasety zverejnené na inom webovom sídle ako data.gov.sk a sprístupnené na portáli data.gov.sk cez hypertextový odkaz by mali byť kategorizované osobitne, nie pod formátmi, pretože ide o metadáta k datasetu, nie o samotné dáta datasetu.

¹⁰ <https://www.data.gov/developers/blog/primer-machine-readability-online-documents-and-data>

1.5 Aktualizácia datasetu

V roku 2019 sme analyzovali pravidelnosť aktualizácií datasetov. Ale v roku 2020 nemáme dostupné ľahko spracovateľné metadáta o pravidelnosti aktualizácie datasetov, preto neuvádzame taký prehľad. Uprednostňujeme štatistiky portálu data.gov.sk pre túto analýzu kvôli jednoduchosti a transparentnosti.

Podiel datasetov s aktualizáciou sme zistili podľa dátumu aktualizácie dátového zdroja uvedeného v zozname datasetov od prevádzkovateľa (NASSES)¹¹. Daný dataset môže byť v zozname uvedený viackrát, ak je zverejnený za rôzne zdrojové dáta. Keďže dátum aktualizácie je uvedený pri dátovom zdroji, započítaváme každý dataset osobitne za každé zdrojové dáta. Celkový počet datasetov sa potom líši od predošlých kapitol, uvádzame percentuálny podiel v grafe.



Graf 5: Podiel datasetov s aktualizovaným dátovým zdrojom

Ak má dataset aktualizovaný dátový zdroj, tak v priemere je aktualizovaný za 134 dní. To sme zistili z rozdielu dátumu aktualizácie dátového zdroja a dátumu zverejnenia dátového zdroja ku datasetu.

1.6 Veľkosť datasetu

V roku 2019 sme analyzovali aj veľkosť datasetu. Ale v roku 2020 nemáme dostupné metadáta o veľkosti datasetov, neuvádzame takú štatistiku. V roku 2019 najväčší podiel mali datasety s veľkosťou rádovo 10 MiB.

1.7 Metadáta k datasetu

V nasledujúcich analýzach by sa mohli (automaticky) spracúvať ďalšie metadáta datasetov, dostupné cez API alebo cez webovú stránku datasetu s lepšou sémantikou, napr.:

- Lokalita (geografická lokácia).
- Pravidelnosť aktualizácie a aktuálnosť datasetu.
- Počet referencií na iné datasety.
- Počet spätných referencií z iných datasetov.
- Počet stiahnutí datasetu anonymne.
- Počet pravidelných stiahnutí datasetu cez API.

¹¹ Súbor: modExport_12052020.xlsx (12.05.2020)

1.8 Používatelia datasetov

Na portáli data.gov.sk v záložke Štatistiky sa uvádza počet registrovaných používateľov portálu 4 520, čo je zvýšenie o 19 % oproti roku 2019. Keďže datasety na portáli sú verejne prístupné, používateľ nemusí byť registrovaný, aby ich mohol prezerat'. Registrovaný používateľ dostáva navyše napríklad notifikácie o datasetoch (nových, aktualizovaných, zmazaných a podobe), metadáta k datasetom. Oprávnený registrovaný používateľ môže zverejňovať datasety.

Portál data.gov.sk sprístupňuje služby priamo, alebo cez *Ústredný portál verejnej správy* (www.slovensko.sk), pomocou Katalógu služieb (Nájsť službu). Portál data.gov.sk informuje o týchto poskytovaných službách¹²:

- Poskytovanie dát v otvorených formátoch.
- Poskytovanie metadát k poskytovaným dátam alebo dátovému zdroju.
- Grafická vizualizácia poskytovaných dát.
- Geografická prezentácia poskytovaných dát.
- Vyhľadávanie dát podľa geografickej lokácie.

Ďalej je možné vyhľadávať dáta podľa rôznych kritérií:

- Informácia o najnovších datasetoch.
- Vyhľadávanie datasetov podľa tagov (kľúčových slov).
- Vyhľadávanie datasetov podľa organizácie (poskytovateľa).
- Vyhľadávanie datasetov podľa formátu.
- Vyhľadávanie datasetov podľa licencie.
- Zoradenie vyhľadaných datasetov podľa dátumu alebo názvu.

Verejnosť môže využiť nasledovné podania:

- Žiadosť o sprístupnenie informácií na účely opakovaného použitia podľa zákona o slobodnom prístupe k informáciám.¹³
- Návrh na zverejnenie údajov na Portáli otvorených dát.
- Žiadosť o registráciu aplikácie na Portáli otvorených dát.
- Podnet na úpravu údajov zverejnených na Portáli otvorených dát.
- Žiadosť o zvýšenie kapacity na uloženie datasetov na Portáli otvorených dát (neposkytuje sa).
- Žiadosť o pridelenie zvýhodneného prístupu k údajom zverejneným na Portáli otvorených dát (neposkytuje sa).

¹² <https://data.gov.sk/published-services>

¹³ Zákon č. 211/2000 Z. z., <https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2000/211/>

2 Obsah zverejňovaných datasetov

2.1 Publikačné minimum

Pre zefektívnenie využívania datasetov a API je vhodné zamerať sa na ciele výber povinne zverejňovaných datasetov ústredných orgánov štátnej správy. Vytvoreniu publikačného minima datasetov predchádza prieskum dopytu verejnosti po najžiadanejších datasetoch a API (každoročne od roku 2017). Medzi povinne zverejňované datasety by zároveň mali patriť kategórie otvorených dát sprístupňovaných na iných významných portáloch otvorených dát. Aktuálny **Akčný plán Iniciatívy pre otvorené vládnutie v Slovenskej republike na roky 2020 – 2021**, schválený Uznesením vlády SR č. 553/2019, v kapitole 1 *Otvorené informácie* obsahuje úlohu 1.2 o zverejňovaní datasetov, ktorá vytýčila konkrétne míľniky pre publikačné minimum:

3. *Participatívnym spôsobom vypracovať publikačné minimum pre štátnu správu. Do 30.4.2020.*
4. *Zaviesť publikačné minimum pre štátnu správu do praxe. Termín: do 31.12.2020.*
5. *Po zavedení publikačného minima pre štátnu správu každoročne zverejňovať dáta z publikačného minima pre štátnu správu na portáli otvorených dát data.gov.sk. Termín: od 31.3.2021 a potom každoročne.*

Publikačné minimum zatiaľ nebolo vypracované a zverejnené. Uskutočnili sa diskusie, je niekoľko návrhov alebo vzorov, aký obsah by publikačné minimum malo mať, či už pre ústredné orgány štátnej správy, verejnú správu všeobecne, alebo pre územnú samosprávu, napr.:

*Chýbajúce open data minimum v SR: Príklady zo zahraničia ako začiatok diskusie*¹⁴

*Návrh Open Data Publikačného Minima pre samosprávy: Zapojte sa do diskusie*¹⁵

*Transparentná samospráva*¹⁶

*Vzorové publikační plány (ČR)*¹⁷

Smernica EP č. 2019/1024¹⁸ o otvorených dátach a opakovanom použití informácií verejného sektora vymedzuje oblasti datasetov s vysokou hodnotou: Geopriestorové údaje; Pozorovanie Zeme a životné prostredie; Meteorológia; Štatistika; Spoločnosti a vlastníctvo spoločností; Mobilita.

Portál európskych údajov (europeandataportal.eu)¹⁹ obsahuje tieto kategórie datasetov: Spravodlivosť, právny systém a verejná bezpečnosť; Poľnohospodárstvo, rybolov, lesníctvo a potravinárstvo; Životné prostredie; Regióny a mestá; Veda a technológia; Ekonomika a financie; Vláda a verejný sektor; Obyvateľstvo a spoločnosť; Doprava; Vzdelávanie, kultúra a šport; Zdravie; Energetika; Predbežné údaje; Medzinárodné vzťahy.

*Open Knowledge Foundation Global Open Data Index (OKFN GODI)*²⁰ zatrieduje datasety do týchto kategórií: Štátny rozpočet; Národné štatistiky; Verejné obstarávanie; Národná legislatíva;

14 <http://www.alvaria.sk/chybajuca-publikacne-open-data-minimum/> (13.11.2019)

15 <http://www.alvaria.sk/open-data-publikacne-minimum-pre-samospravy/> (18.06.2020)

16 <https://utopia.sk/wiki/display/opendata/Transparentna+samosprava> (30.4.2019)

17 <https://opendata.gov.cz/vzor:start>

18 https://datalab.digital/wp-content/uploads/Smernica2019_1024OPENDATA.pdf

19 <https://www.europeandataportal.eu/data/datasets?locale=sk>

20 <https://index.okfn.org/dataset/>

Katastrálne údaje; Legislatívne návrhy; Kvalita ovzdušia; Národné mapy; Obchodný register; Výsledky volieb; Lokality; Kvalita vody; Vládne výdavky; Vlastníctvo pôdy.

2.2 Prieskum dopytu verejnosti po najžiadanejších datasetoch a API

Pri zverejňovaní datasetov ústredných orgánov štátnej správy sú zverejňované aj datasety s nízkou kvalitou a najmä nízkou užitočnosťou. Medzi niektoré dôvody môžu patriť nedostatočná (formálna) kontrola zverejňovania datasetov, či nestotožnenie sa s filozofiou zverejňovania dát, prípadne nízke povedomie o využiteľnosti otvorených dát zo strany verejnej správy aj občianskej spoločnosti či súkromného sektora.

Úrad splnomocnenca vlády SR pre rozvoj občianskej spoločnosti každoročne od roku 2017 vykonáva prieskum dopytu verejnosti po najžiadanejších datasetoch a API.²¹ Tieto datasety by mali byť zverejňované prioritne. Predpokladáme, že užitočnosť verejnosťou žiadaných datasetov je objektívne vysoká. Verejnosť žiada datasety všeobecne od verejnej správy aj územnej samosprávy. Občan alebo podnikateľ nemusí vedieť, ktorá organizácia je vlastníkom požadovaných dát. V tejto analýze skúmame zverejňovanie datasetov najmä od ústredných orgánov štátnej správy.

Počas januára a februára 2020 sa v prieskume zisťoval dopyt po datasetoch, ktoré priamo súvisia s princípmi otvoreného vládnutia (transparentnosť, participácia).

Tabuľka 2: Najžiadanejšie datasety súvisiace s otvoreným vládnutím²²

| |
|--|
| Dáta o trvaní jednotlivých procesných úkonov v trestnom konaní |
| Dáta o trestnej činnosti na úrovni ulíc |
| Dáta týkajúce sa firiem s majetkovou účasťou štátu a ďalších verejných inštitúcií |
| Zoznam rozpočtových, príspevkových a neziskových organizácií verejnej správy so zriaďovateľmi |
| Centrálne evidencie majetku, register ponúkaného majetku štátu |
| Zmluvy, objednávky, faktúry rezortov |
| Register výberových konaní |
| Počty zamestnancov v štátnej správe |
| Všetky referenčné registre |
| Register právnických osôb so všetkými integrálnymi registrami |
| Register politických strán a politických hnutí |
| Dotácie rezortov (okrem informácií, ktoré sa zverejňujú cez modul dotačných schém) |
| Dáta z informačných systémov Dátového centra obcí a miest (DCOM), ktoré spravuje DEUS, vrátane rozpočtových údajov, registra adries, počtu obyvateľov, evidencie psov, zoznamu daňových dlžníkov, zoznamu obchodných prevádzok, zoznamu žiadostí a podaní od občanov, zoznamu nehnuteľností v správe samospráv či informácií o územnom plánovaní |
| Dáta o znečistení ovzdušia, vôd či rezíduá pesticídov v potravinách |

21 Pravidelné úlohy B.8 a B.9 z Akčného plánu Iniciatívy pre otvorené vládnutie v Slovenskej republike na roky 2017 – 2019, schváleného Uznesením vlády SR č. 104/2017.

22 https://www.minv.sk/?ros_vsetky-spravy&sprava=prieskum-dopytu-verejnosti-po-najzidanejsich-datasetoch-a-api

Úrad vlády SR potvrdil plánovanú aktualizáciu Centrálného registra zmlúv (crz.gov.sk), ktorý bol súčasťou podnetov na zverejňovanie informácií v forme otvorených dát.

Výsledkom prieskumu dopytu verejnosti po najžiadanejších datasetoch a API je zverejnená krátka správa²³ a tabuľka²⁴, ktorá je príliš veľká na zobrazenie vcelku na jednej strane. Zistené dopyty sumarizujeme po skupinách v nasledujúcich tabuľkách.

Tabuľka 3: Najžiadanejšie datasety – zverejnené

| Organizácia | Stručný popis dát (sumarizácia) |
|---|---|
| Ministerstvo dopravy a výstavby SR | Zoznam organizácií, otváracie hodiny, počet zamestnancov |
| | Zoznam železničných staníc a zastávok |
| | Mapy cestnej siete, plánovaná dopravná infraštruktúra |
| | Intenzita a skladba dopravy – mýtny systém NDS |
| | Centrálna technická evidencia ciest – kvalita vozoviek |
| Ministerstvo pôdohospodárstva a rozvoja vidieka SR | Pôda – mapa, pozemkové úpravy, pôdne jednotky, pôdne bloky, trhovacia cena, analýza pôdných vzoriek na živiny, pH pôdy, lesné hospodárske celky, lesné porasty, poľovné revíry, vinohrady, zavlažovanie, správa |
| | Hnojivá, krmivá, sekundárne zdroje živín, chemická ochrana rastlín |
| | Genetické zdroje rastlín, produkcia plodín, včelstvá, rybolov |
| | Certifikáty, osvedčenia, registre, zmluvy, faktúry, objednávky, cenníky |
| | Projekty – hodnotitelia, stav žiadostí, príspevky, manažéri, kontrolóri |
| | Potravinárske, živočíšne, veterinárne prevádzky, aplikačné zariadenia |
| | Laboratória |
| | Vzájomne uznané výrobky EÚ |
| Databáza filmov Agrofilm | |
| Ministerstvo kultúry SR | Register národných kultúrnych pamiatok |
| Protimonopolný úrad SR | Dáta o kontrolách |
| Úrad geodézie, kartografie a katastra SR | Ortofotomozaika |
| | Zoznam vrcholov (geografické názvoslovie) |
| Úrad podpredsedu vlády SR pre investície a informatizáciu | Dáta o eurofondoch a projektoch 2014 – 2020, ITMS2014+ |
| | Dáta o eurofondoch a projektoch 2007 – 2013 |
| Úrad pre verejné obstarávanie | Vestník verejného obstarávania |
| Úrad vlády SR | Zoznam subjektov s aktivovanou elektronickou schránkou |

23 https://www.minv.sk/?ros_vsetky-spravy&sprava=prieskum-dopytu-verejnosti-po-najzidanejsich-datasetoch-a-api

24 <https://docs.google.com/spreadsheets/d/13uVKtcS4vDHZISeLZvv-0E3Cefmutc8P-CVk0cEf1RM/edit?usp=sharing>

Tabuľka 4: Najžiadanejšie datasety – čiastočne zverejnené

| Organizácia | Stručný popis dát (sumarizácia) |
|--|--|
| Ministerstvo dopravy a výstavby SR | Cestovné poriadky |
| | Zoznamy štátnych budov |
| | Register homologizovaných typov vozidiel |
| Ministerstvo spravodlivosti SR | Rozhodnutia (obcan.justice.sk), exekúcie |
| | Priemerná dĺžka konania, výkonnosť súdnictva |
| Ministerstvo školstva, vedy, výskumu a športu SR | Repozitár otvorených vzdelávacích zdrojov |
| | Financovanie školstva, granty |
| | Anonymizované údaje z centrálnych registrov rezortu školstva |
| | Zoznam škôl, študijných/učebných odborov, výsledky testovania žiakov |
| Ministerstvo vnútra SR | Štatistiky o kriminalite, priestupkoch a pokutách (s lokalitou) |
| Ministerstvo životného prostredia SR | Atmosférické podmienky, meteorologické grafické prvky (INSPIRE) |
| | Dáta o životnom prostredí – ovzdušie, voda, environmentálne záťaž |
| | Zoznam prírodných pamiatok, chránených území, rieky (INSPIRE) |
| Štatistický úrad SR | Register právnických osôb, register ekonomických subjektov |

Tabuľka 5: Najžiadanejšie datasety – nezverejnené

| Organizácia | Stručný popis dát (sumarizácia) |
|---|--|
| Ministerstvo dopravy a výstavby SR | Polohy a informácie o vlakoch; vyťaženosť vlakov, staníc, zastávok |
| | Zoznam autobusových staníc a zastávok; vyťaženosť autobusov, staníc |
| | Register prihlásených vozidiel |
| | Inžinierske siete (geodáta) |
| Ministerstvo financií SR | Rozpočtový informačný systém (RIS) štátnej správy, verejnej správy, štátnych podnikov, územnej samosprávy, aj kontrola a audit |
| | Párovanie IČO, DIČ, IČ DPH, fakturačné údaje, účtovné závierky |
| | Anonymizované daňové priznania FO |
| | Centrálna evidencia majetku, register ponúkaného majetku štátu |
| Ministerstvo hospodárstva SR | Územné plány, zdroje energie, ložiská nerastných surovín |
| | Zoznam nahlásených nevyhovujúcich potravín, prevádzok |
| | Výrobné a priemyselné zariadenia, ubytovacie zariadenia |
| Ministerstvo práce, sociálnych vecí a rodiny SR | Počet uchádzačov o zamestnanie za každý mesiac na úrovni úradov práce |

| | |
|--|---|
| Ministerstvo spravodlivosti SR | Údaje z obchodného registra (pozri: Register právnických osôb) |
| | Právne predpisy a legislatívny proces v strojovo čitateľnej forme |
| Ministerstvo školstva, vedy, výskumu a športu SR | Dáta z informačného systému športu |
| | IUVENTA – zoznam školení |
| Ministerstvo vnútra SR | Živnostenský register SR (pozri: Register právnických osôb) |
| | Register prihlásených vozidiel |
| | Mapa pokrytia verejného priestranstva kamerami |
| | Register spoločenstiev vlastníkov bytov a nebytových priestorov |
| | Register politických strán a politických hnutí |
| | Adresné body (geodáta) |
| Ministerstvo zdravotníctva SR | Zmluvy VŠZP s poskytovateľmi zdravotnej starostlivosti |
| | Údaje o vykázanej zdravotnej starostlivosti od zdravotných poisťovní pre NCZI na úrovni spravodajskej jednotky |
| | Ročný a štvrťročný výkaz o ekonomike organizácií v zdravotníctve, o počte a štruktúre pracovníkov |
| | Štvrťročné výkazy L1 a L2 o spotrebe liekov |
| | Anonymizované údaje o čakacích lehotách za pacienta a WL výkon |
| Najvyšší kontrolný úrad SR | Dáta o kontrolách |
| Národná banka Slovenska | Štatistiky |
| Štatistický úrad SR | Počty zamestnancov v štátnej správe |
| Úrad geodézie, kartografie a katastra SR | Kataster – číselníky, katastrálne územia, listy vlastníctva, katastrálne konania, pozemky, stavby, budovy, byty |
| | ZBGIS – základná báza údajov pre geografický informačný systém |
| | Digitálny model reliéfu, digitálny model terénu, výškový model |
| Úrad pre verejné obstarávanie | Referencie verejného obstarávania, profily, kontroly |
| | Protokoly z elektronických aukcií |
| | Register konečných užívateľov výhod |
| Úrad vlády SR | Majetkové priznania verejných funkcionárov, s históriou |
| | Plnenia úloh z uznesení vlády SR |
| | Centrálny register zmlúv v strojovo čitateľnej forme (crz.gov.sk) |
| | Všetky centrálné číselníky |
| | Register výberových konaní |

Tabuľka 6: Najžiadanejšie API

| Organizácia | Stručný popis dát a API (sumarizácia) |
|---|--|
| Ministerstvo doprav a výstavby SR / územná samospráva | Aktuálna poloha vozidiel MHD, autobusov a vlakov |
| | Dopravná sieť |
| Ministerstvo spravodlivosti SR | Exekúcie |
| | Obchodný register SR |
| | Slov-Lex |
| Ministerstvo vnútra SR | Živnostenský register SR |
| | Dáta o dopravných obmedzeniach a nehodách |
| | Register adries, vyhľadávanie polohy |
| | Register prihlásených vozidiel |
| | Register spoločenstiev vlastníkov bytov a nebytových priestorov |
| Ministerstvo zdravotníctva SR | Výskyt alergénov |
| Ministerstvo životného prostredia SR | Meteorologické údaje (http://meteo.shmu.sk/customer/home/opendata/) |
| | Dáta z predikčných modelov počasia SHMÚ |
| | Dáta z predikčných modelov povodní |
| Úrad geodézie, kartografie a katastra SR | ZBGIS vo vektorovom formáte |
| | Digitálny model reliéfu, digitálny model terénu, výškový model |
| | Ortofotomapa, ortofotomozaika |
| | Kataster nehnuteľností |
| Úrad vlády SR | Autentifikácia obyvateľov |
| | Prihlasovanie k ústrednému portálu verejnej správy www.slovensko.sk |
| | Prístup k elektronickej schránke |

2.3 Správa o implementácii akčného plánu 2017 – 2019

Iniciatíva pre otvorené vládnutie je globálne partnerstvo, ktoré spája vládných reformátorov a lídrov občianskej spoločnosti pri tvorbe akčných plánov, vďaka ktorým budú vlády otvorejšie, prístupnejšie a zodpovednejšie. Závazky akčného plánu môžu vychádzať z existujúcich snáh identifikovať nové kroky na dokončenie už prebiehajúcich reforiem alebo z podnetu iniciovať opatrenia v úplne novej oblasti. Nezávislý hodnotiaci mechanizmus monitoruje všetky akčné plány s cieľom zabezpečiť dodržiavanie a plnenie záväzkov.²⁵

Správa o implementácii akčného plánu 2017 – 2019 popisuje naplnenie záväzkov akčného plánu hodnotených v dvoch špecifických správach (Správa o príprave a tvorbe akčného plánu, Správa o implementácii akčného plánu). V nasledujúcej tabuľke sumarizujeme niektoré závery hodnotenia relevantných záväzkov z kapitoly Otvorené informácie.

Tabuľka 7: Skrátený prehľad hodnotenia relevantných záväzkov akčného plánu 2017 – 2019

| | |
|--|---|
| Záväzok č. 6 | Na základe aktualizovaného zoznamu datasetov v gescii ministerstiev a nimi zriadených organizácií aj ostatných ústredných orgánov štátnej správy zverejniť datasety v súlade s plánom ich zverejnenia a aktualizovania na portáli otvorených dát. |
| Záväzok sa implementuje priebežne, s rôznou mierou dokončenia. Niektoré orgány ústrednej štátnej správy zverejnili veľmi málo datasetov, napríklad Ministerstvo obrany SR, Ministerstvo financií, Ministerstvo práce, sociálnych vecí a rodiny SR. | |
| Záväzok č. 8 | Vykonať prieskum dopytu verejnosti po najžiadanejších datasetoch otvorených dát. |
| Národná agentúra pre sieťové a elektronické služby (NASES) by mohla pravidelne poskytovať štatistiky návštevnosti a analytické dáta o prevádzke data.gov.sk , najmä počet návštevníkov konkrétnych datasetov. | |
| Záväzok č. 9 | Na základe výsledkov prieskumu dopytu verejnosti po najžiadanejších datasetoch otvorených dát zverejniť najžiadanejšie datasety na portáli otvorených dát v súlade s platnou legislatívou. |
| Miera dokončenia tohto záväzku je limitovaná. Keďže najžiadanejšie datasety stále nie sú dostupné verejnosti, tento záväzok k otvoreniu vládnutiu, najmä k prístupu k informáciám, prispel len minimálne. | |
| Záväzok č. 10 | Vypracovať analýzu zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát a predložiť ju na zasadnutie Rady vlády SR pre mimovládne neziskové organizácie. |
| Zmapovanie zverejňovania otvorených dát orgánmi ústrednej štátnej správy je užitočné. Prispieva k otvoreniu vládnutiu, a to poskytovaním podrobných informácií o napredovaní tohto aj predchádzajúcich záväzkov a identifikovaním prekážok, ktoré bránia pokroku. Od prvej analýzy v roku 2017 až doteraz boli však zlepšenia vo zverejňovaní otvorených dát okrajové. | |

25 Nezávislý hodnotiaci mechanizmus (IRM): Slovensko – Správa o implementácii akčného plánu 2017 – 2019 (Mária Žuffová, 2020):
<https://www.opengovpartnership.org/documents/slovakia-implementation-report-2017-2019-for-public-comment/>
https://www.opengovpartnership.org/wp-content/uploads/2020/03/Slovakia_Implementation_Report_2017-2019_SK_for-public-comment.pdf

| | |
|---|--|
| Závazok č. 13 | Zabezpečiť zverejňovanie otvorených dát a sprístupňovanie verejne dostupných aplikačných programových rozhraní v projektoch financovaných prostredníctvom Operačného programu integrovaná infraštruktúra a z iných verejných zdrojov. |
| Závazok je splnený v obmedzenej miere z dôvodu, že vyjadrenie úmyslu sprístupniť API je síce dôležitý prvý krok, ten však nezaručuje, že k sprístupneniu nakoniec dôjde. Nie je jasné, aké bude mať nenaplnenie deklarovaneho zámeru pre inštitúcie verejnej správy dôsledky. Kľúčové informačné systémy verejnej správy v súčasnosti neposkytujú otvorené API. Závazok preto neprispel k otvorenému vládnutiu, nakoľko sa tento stav nezmenil. | |
| Závazok č. 14 | Vykonať prieskum dopytu verejnosti po najžiadanejších verejne dostupných aplikačných programových rozhraniach. |
| Závazok možno považovať za splnený. Okrajovo prispel k otvorenému vládnutiu, pretože jeho splnenie prinieslo nové informácie o dopyte po otvorených API. Pre časť verejnosti so záujmom o otvorené API záväzok navyše vytvoril príležitosť zapojiť sa do procesu a sformulovať svoje preferencie. | |
| Závazok č. 15 | Na základe výsledkov prieskumu dopytu verejnosti po najžiadanejších verejne dostupných aplikačných programových rozhraniach sprístupniť v súlade s platnou legislatívou a existujúcimi technickými podmienkami tie najžiadanejšie spolu s uvedením licencie (vzťahov) na ich využívanie. |
| V správe o stave slovenského e-governmentu, ktorú nedávno zverejnilo Slovensko.Digital, sa konštatuje, že 90 % aktivít súvisiacich so sprístupnením API sa oneskorilo a API, ktoré sú v súčasnosti dostupné z informačných systémov verejnej správy, sú neuspokojivé a nespĺňajú štandardy. Keďže bol tento záväzok splnený len sčasti, jeho príspevok k otvorenému vládnutiu bol okrajový. Na druhej strane je potrebné uviesť, že oslovený zástupca Slovensko.Digital tvrdí, že oficiálny záväzok vlády sprístupňovať API im ako občianskemu združeniu uľahčuje komunikáciu s orgánmi ústrednej štátnej správy v tejto oblasti. | |

2.4 Užitočnosť otvorených dát pre používateľa

V článku *Towards a Data Quality Score in open data*²⁶ sa na príklade Toronta (Kanada) zrozumiteľne opisuje pohľad používateľa pri využívaní otvorených dát. Hlavná myšlienka sa dá zhrnúť do jednej vety: **Kvalita je dôležitejšia ako kvantita**. Kvalita má niekoľko charakteristík, s percentuálnom zastúpením v celkovom hodnotení kvality:

1. Použiteľnosť (ako ľahko sa pracuje s dátami) 38 %
2. Metadáta (ako sú dáta vysvetlené) 25 %
3. Aktuálnosť (aká je pravidelnosť aktualizácie dát) 18 %
4. Úplnosť (či sú dáta kompletné alebo obsahujú prázdne hodnoty) 12 %
5. Prístupnosť (či sa dá použiť API) 7 %

26 <https://medium.com/open-data-toronto/towards-a-data-quality-score-in-open-data-part-1-525e59f729e9>

3 Kvalitatívna analýza

Táto kapitola analyzuje zverejňovanie datasetov z pohľadu používateľa, podľa kvalitatívnych kritérií, napr. aké nástroje a aplikácie môže používateľ použiť, alebo v porovnaní so zahraničím.

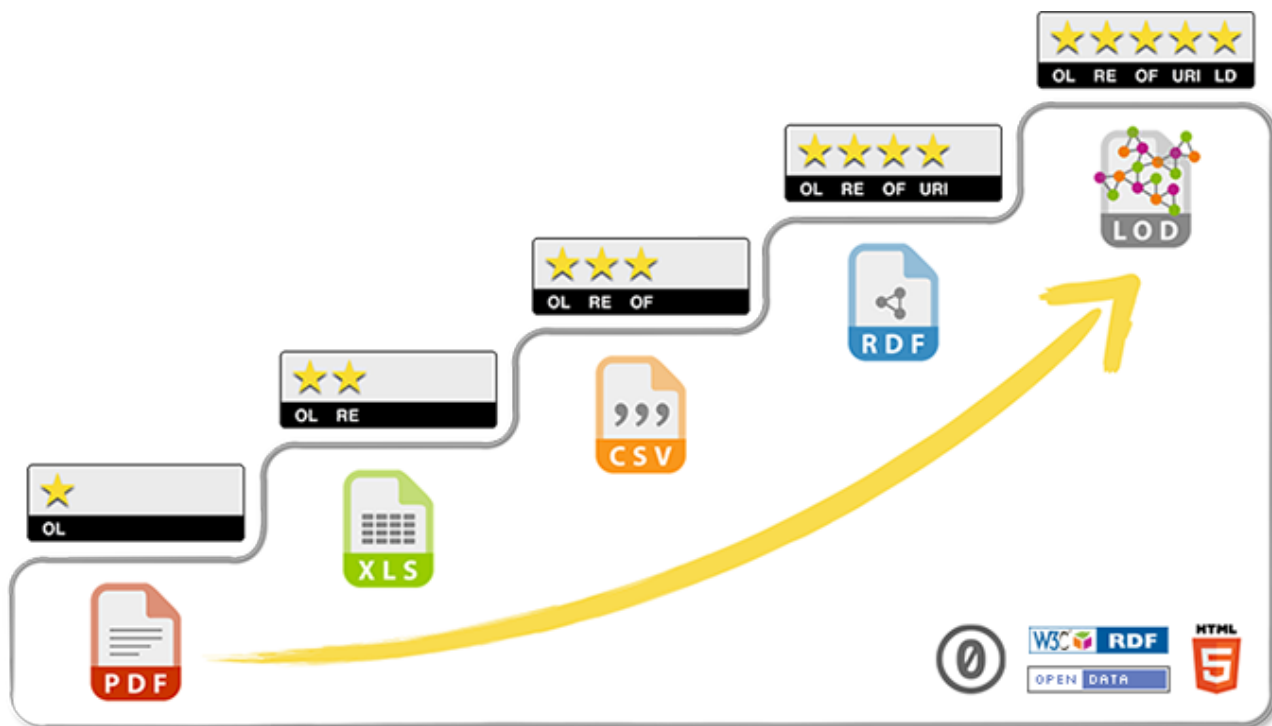
Špeciálnym typom datasetu je katalóg (zoznam) zverejňovaných datasetov danej organizácie.

3.1 Hodnotenie datasetov

Tim Berners-Lee, vynálezca webu a iniciátor prepojených dát, navrhol hodnotenie otvorených dát hviezdikami²⁷:

- * zverejnené dáta s otvorenou (verejnou) licenciou (vo webovom prostredí)
- ** + štruktúrované strojovo spracovateľné dáta
- *** + otvorený štandardizovaný formát
- **** + permanentná URI adresa (permanentný referencovateľný identifikátor)
- ***** + prepojené dáta (prepojené cez referencovateľný identifikátor na iné dáta)

Kvalita poskytovaného datasetu je v tomto význame zadefinovaná vo vyhláške ÚPVII č. 78/2020 Z. z. (Vyhláška o štandardoch pre informačné technológie verejnej správy)²⁸, § 38 až § 40.



Obrázok 2: Hodnotenie otvorených dát hviezdikami

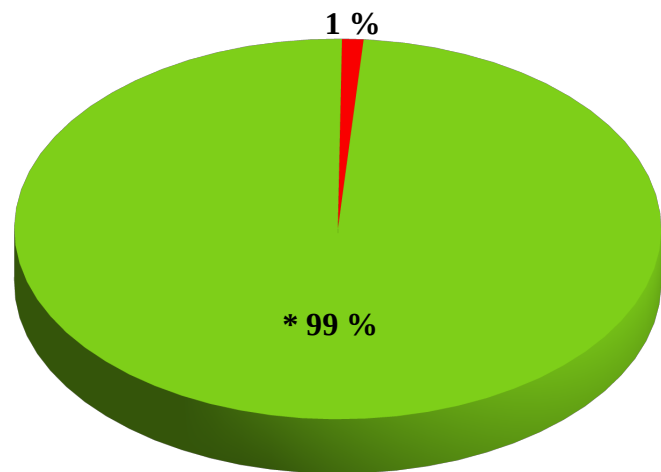
Hodnotenie hviezdikami možno pridať k vyššie uvedeným prehľadom licencií (kap. 1.3), formátov (kap. 1.4). Minimálna požadovaná úroveň pre otvorené dáta je úroveň 3 hviezdíček.

²⁷ <https://5stardata.info/cs/>

²⁸ <https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2020/78/20200501>

Ohodnotenie výberu licencie hviezdíčkami je jednoduché. Všetky typy licencií uvedené na portáli data.gov.sk, ktoré boli použité v grafe 3 (licencie datasetov), sme rozdelili na dve skupiny – vhodná a nevhodná. Vhodné alebo štandardmi odporúčané sú (otvorené) verejné licencie CC0, CC-BY, CC-BY-SA, GNU FDL a pod. Taká licencia má jednu hviezdíčku. Nevhodná licencia má nula hviezdíček.

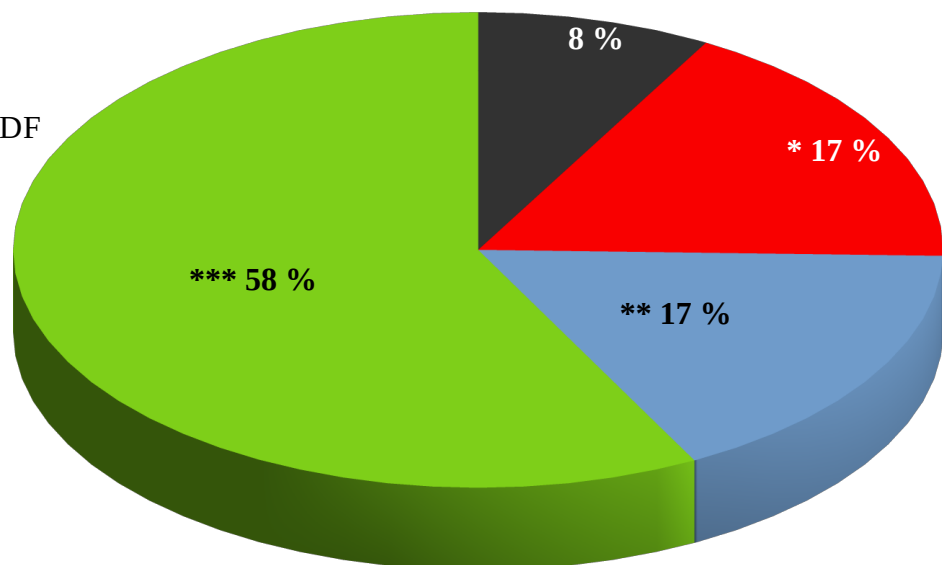
- vhodná
- nevhodná



Graf 6: Hodnotenie licencie hviezdíčkami

Ďalej sme ohodnotili formáty zverejnených datasetov hviezdíčkami podľa dát pre graf 4 (formáty datasetov). Ľahko sa dajú určiť hodnotenia jednou, dvomi, tromi hviezdíčkami.

- *** CSV/XML/JSON
- ** XLSX/XLS
- * HTML/XHTML/PDF
- iné (nezaradené)



Graf 7: Hodnotenie formátu hviezdíčkami

V tomto roku nemáme dostupné ľahko spracovateľné dáta o jednotlivých datasetoch také, aby sa dal urobiť prienik hodnotenia licencií datasetov a formátov datasetov.

Formát RDF, ktorý poskytuje ontológiu pre metadáta ku datasetu, je na portáli data.gov.sk len pri 6 datasetoch, čo je zanedbateľne málo. Iné formáty, napr. JSON-LD, s podobným účelom sa môžu na portáli vyskytovať (takisto v malom počte).

Podmienkou pre dataset úrovne 4 hviezdíček je použitie ontológie pri definícii údajov (metadátach), použitie referencovateľného identifikátora (permanentnej URI/URL) pre dataset a referencovateľných identifikátorov v dátach. Táto problematika nie je vo zverejnených datasetoch dostatočne dobre hodnotiteľná.

Úroveň 5 hviezdíček spĺňajú tie datasety, ktoré majú 4 hviezdíčky a obsahujú v dátach prepojenia referencovateľnými identifikátormi na iné datasety. Toto nevieme dostatočne rýchlo vyhodnotiť z dostupných informácií o datasetoch.

Séria šiestich datasetov od Ministerstva životného prostredia SR alebo jeho podriadených organizácií je dobrým príkladom 5-hviezdičkových datasetov. Sú zverejnené s licenciou CC-BY, vo formáte RDF a sú prepojené, čo je navyše indikované tagom *Linked Data*.

3.2 Global open data index

<https://index.okfn.org>

Slovensko sa umiestnilo v poslednom sledovanom období 2016-2017 na 32. mieste so 47 % v rebríčku 94 sledovaných štátov (Taiwan na 1. mieste s 90 %). Metodika je založená na technickej a právnej otvorenosti vybraných datasetov v oblastiach: štátny rozpočet, národné štatistiky, verejné obstarávanie, národná legislatíva, katastrálne údaje, legislatívne návrhy, kvalita ovzdušia, národné mapy, obchodný register, výsledky volieb, lokality, kvalita vody, vládne výdavky, vlastníctvo pôdy. Slovensko zaznamenalo zlepšenie oproti roku 2015, kedy bolo na 50. mieste s 35 %. Slovensko má výborné hodnotenie z národných štatistík, verejného obstarávania, lokalít; zlé hodnotenie z národných máp, výsledkov volieb, kvality vody, vládnych výdavkov.

3.3 Open data barometer

<https://opendatabarometer.org>

Najnovšie sledované obdobie je za rok 2017. Slovensko nie je zaradené do rebríčka v tomto období. V predchádzajúcom období 2016 sa Slovensko umiestnilo na 30. mieste s 45 bodmi zo 100 v rebríčku 92 štátov (Spojené kráľovstvo na 1. mieste so 100 bodmi). Oproti roku 2015 je to zlepšenie o 7 bodov. Skúmané sú tri kategórie kritérií (body pre Slovensko 2016): pripravenosť (59), implementácia (43), prínos (34).

3.4 Open data watch

<https://opendatawatch.com>

Slovensko sa umiestnilo v poslednom sledovanom období 2018²⁹ na 12. mieste so 76 % v rebríčku 178 hodnotených krajín (Singapur na 1. mieste s 86 %). Oproti predošlému obdobiu 2017 síce kleslo o 1 miesto, ale zlepšilo si percentuálne hodnotenie. Posudzujú sa oficiálne zverejňované štatistiky štátu v oblastiach: sociálna, ekonomická, životné prostredie. Prehľad o Slovensku za rok 2018³⁰ je na základe dát Štatistického úradu SR³¹.

29 <http://odin.opendatawatch.com/Report/rankings?sortOrder=&appConfigId=5>

30 <http://odin.opendatawatch.com/Report/countryProfile/SVK?appConfigId=5>

31 <http://www.statistics.sk>

3.5 Index digitálnej ekonomiky a spoločnosti (DESI)

<https://ec.europa.eu/digital-single-market/en/desi>

V indexe digitálnej ekonomiky a spoločnosti (DESI) 2019³² patrí Slovensku 21. miesto so 46,3 bodmi zo 100 v rebríčku 28 krajín EÚ, s poklesom o 1 miesto oproti roku 2018 (pri zvýšení bodov). Slovensku sa darí v ukazovateli otvorených dát, kde dosiahlo 8. miesto (nad priemerom EÚ)³³. Index DESI hodnotí členské štáty EÚ vo výkonnosti podľa stratégie pre jednotný digitálny trh a digitálnu konkurencieschopnosť.

3.6 Open data maturity index

<https://www.europeandataportal.eu>

Slovensko sa umiestnilo v poslednom sledovanom období 2019³⁴ na 28. mieste s 33 % v rebríčku 31 sledovaných európskych štátov (EÚ28+), kde Írsko je na 1. mieste s 90 %. Metodika je založená na posúdení 4 dimenzií otvorených dát: politika, portál, dopad, kvalita.

Slovensko zaznamenalo prudký prepád oproti roku 2018 (9. miesto so 74 %). V roku 2019 je Slovensko hlboko pod priemerom, zaradené do poslednej skupiny začiatočníkov (lídri trendu, rýchli nasledovníci trendu, nasledovníci, začiatočníci). Podrobnejší pohľad na Slovensko 2019³⁵ ukazuje, že prepád nastal v roku 2018 a prehĺbil sa v roku 2019. Najhorší výsledok je v dimenzii dopad, najmä v kritériách: povedomie, spoločnosť, životné prostredie, ekonomika. Dimenzia politika má zlé hodnotenie podľa kritéria implementácia. Dimenzia portál má zlé hodnotenie podľa kritérií poskytovanie dát a udržateľnosť portálu.

Bariéry pre Slovensko v otvorených dátach: relatívne malá dostupnosť datasetov s vysokou kvalitou, obmedzené chápanie výhod opakovaného použitia otvorených dát, strach z porušenia ochrany osobných údajov (GDPR).

3.7 Nástroje

Na portáli data.gov.sk v záložke Nástroje je dostupný iba jeden nástroj (žiadna zmena oproti roku 2019). Vraj by mal byť dostupný niekde ešte jeden nástroj. Na tomto alebo podobnom mieste by mohli byť zverejňované ďalšie nástroje a API pre prácu s datasetmi, vo forme softvéru s otvoreným zdrojovým kódom a s verejnou licenciou EUPL.

SPARQL editor

Tento nástroj slúži na vytváranie RDF dopytov pre SPARQL (*SPARQL protocol and RDF query language*) a prezentáciu získaných dát priamo vo webovom prehliadači. SPARQL môže byť použitý na prácu s dátami a metadátami vo formáte RDF (*resource description framework*). Vzhľadom na veľmi malý počet datasetov s formátom RDF, je tento nástroj nevyužitý.

32 https://ec.europa.eu/slovakia/news/desi2019_sk

33 https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60005

34 <https://www.europeandataportal.eu/en/highlights/open-data-maturity-report-2019-out-now>,
https://www.europeandataportal.eu/sites/default/files/open_data_maturity_report_2019.pdf

35 https://www.europeandataportal.eu/sites/default/files/country-factsheet_slovakia_2019.pdf

3.8 Aplikácie

Na portáli data.gov.sk v záložke Aplikácie sú dostupné iba 3 aplikácie (pokles o 1 oproti roku 2019). Na tomto mieste by mali byť zverejnené rôzne používateľské aplikácie na prácu so zverejnenými dátami (datasetmi). Prekážkou pri zverejňovaní aplikácií je povinné použitie eID, kvôli problémom so samotným použitím eID.

Aktuálne údaje centrálnej technickej evidencie ciest (Mapy CDB)

Aplikácia sprístupňuje aktuálne údaje centrálnej technickej evidencie ciest.

Slovenská správa ciest, 2019, <https://ismcs.cdb.sk/portal/mapviewer>

Notifikátor výberových konaní

E-mailové notifikácie pre novo-vyhlásené výberové konania z portálu <https://open.slovensko.sk>.

Projekt Výsledky a štatistiky verejného obstarávania a vizualizácia dát, 2017,

<https://www.uvostat.sk/vyberovekonania>

OpenSpending mesto Prešov

Aplikácia na vyhľadávanie, vizualizáciu a analýzu rozpočtových údajov miest (Prešov a iných).

Open knowledge foundation, 2017, <https://openspending.org/viewer/1ae301d0fb4c5bd535ce8e2eadf7a668:presov.sk-program-budget-2016?lang=en>

3.9 API

Portál data.gov.sk poskytuje CKAN API³⁶ pre prácu s datasetmi a metadátami datasetov. API³⁷ môže slúžiť na automatizované spracovanie dát a metadát pre používateľov, aj pre potreby budúcich analýz zverejňovania datasetov.

Dokumentácia CKAN (*Comprehensive Knowledge Archive Network*)³⁸ je dostupná už v novej verzii.

36 <https://data.gov.sk/api>

37 https://sk.wikipedia.org/wiki/Application_programming_interface

38 <https://docs.ckan.org/en/2.8/>
<https://en.wikipedia.org/wiki/CKAN>
<https://github.com/ckan/ckan>

Záver

Podstatné zistenia z analýzy uvádzame na záver vo forme **odporúčaní na najbližšie obdobie:**

- **Prieskum dopytu verejnosti po najžiadanejších datasetoch a API** – pokračovať v prieskume priebežne, s vyhodnotením ročne.
- **Monitorovať a vyhodnocovať zverejnenie dopytu verejnosti** po najžiadanejších datasetoch orgánov územnej samosprávy.
- **Publikačné minimum** – participatívnym spôsobom vypracovať publikačné minimum pre štátnu správu (najlepšie s rozšírením na celú verejnú správu a územnú samosprávu), následne ho zaviesť a podľa neho zverejňovať datasety.
- **Zvýšiť kvalitu a užitočnosť datasetov** – portál data.gov.sk sa už niekoľko rokov postupne plní datasetmi. V nasledujúcich rokoch by sme sa mali viac zaujímať o dostupnosť datasetov s vysokou kvalitou. To je jedna z identifikovaných bariér *Open data maturity index 2019* pre Slovensko (kap. 3.6).
- **Rozšírenie záberu dát** – vzhľadom na hodnotenie otvorených dát hviezdikami (kap. 3.1), na *Stratégiu digitálnej transformácie Slovenska 2030*³⁹, na pripravovaný program *Digitálna Európa 2021 – 2027*⁴⁰ a na *Európsku dátovú stratégiu (EK 19.02.2020)*, je vhodné, aby sa zvýšená pozornosť venovala týmto témam: datasety samosprávy (obce, mestá, kraje), otvorené prepojené dáta (interoperabilita), otvorené veľké dáta, umelá inteligencia, GDPR, kybernetická bezpečnosť.
- **Otvorené API, otvorené štandardy** – cieľom zverejňovania otvorených dát ústredných orgánov štátnej správy (aj celej verejnej správy) je zvyšovanie transparentnosti a podpora hospodárstva založeného na dátach. K tomu potrebujeme otvorené API, otvorené štandardy. Preto je vhodné, aby sa ďalej podporovali témy: otvorené API, otvorené štandardy pre interoperabilitu, moje dáta a interoperabilita, umelá inteligencia, inteligentné mesto a internet vecí, otvorený softvér a otvorený hardvér.
- **Zvýšiť povedomie o výhodách opakovaného použitia otvorených dát** – napríklad väčším prepojením verejných politík v oblastiach digitalizácia, vzdelávanie, podnikanie. To je jedna z identifikovaných bariér *Open data maturity index 2019* pre Slovensko (kap. 3.6).
- **Podporiť zverejnenie existujúcich aplikácií a vznik nových** – na portáli data.gov.sk v záložke Aplikácie by malo byť podstatne viac zverejnených aplikácií spracúvajúcich otvorené dáta, aby sa dosiahol pozitívny dopad na spoločnosť a ekonomiku.
- V ďalšom *Akčnom pláne Iniciatívy pre otvorené vládnutie v Slovenskej republike* by mali byť témy o otvorených dátach previazané aj na kapitolu **Otvorené vzdelávanie a otvorená veda**.

39 <https://www.vicpremier.gov.sk/sekcie/informatizacia/digitalna-transformacia/strategia-digitalnej-transformacie-slovenska-2030/index.html>,
<https://www.vicpremier.gov.sk/wp-content/uploads/2019/06/Strategia-digitalnej-transformacie-Slovenska-2030.pdf>

40 EC(2018) 434 final, <https://www.slov-lex.sk/legislativne-procesy/SK/LP/2018/687>

Odporúčania pre prevádzkovateľa portálu data.gov.sk (NASES)

Pri analýze zverejňovania datasetov na portáli otvorených dát (data.gov.sk), ktorého NASES je prevádzkovateľom, sme zaznamenali niekoľko problémov alebo nedostatkov. Väčšina z nich bola zistená už v analýze z roku 2019. Pridávame odporúčania, ako ich vyriešiť:

1. Na portáli data.gov.sk v záložke Štatistiky na obrázku sa sumárne počty datasetov nezhodujú so skutočným počtom datasetov po kliknutí na daný riadok v obrázku. Mal by sa zobrazíť všade aktuálny stav datasetov.
2. Je užitočné, že portál data.gov.sk zverejňuje datasety od každej organizácie zvlášť. Avšak bolo by prehľadnejšie, ak by portál data.gov.sk umožňoval zobrazíť zoznam ústredných orgánov štátnej správy⁴¹ a k nim priradené podriadené organizácie⁴². Napríklad Regionálne úrady verejného zdravotníctva priradené k Ministerstvu zdravotníctva SR. Tiež všetky územné samosprávy združiť do jedného zoznamu.
3. Portál data.gov.sk zobrazuje počty datasetov podľa organizácií spravidla agregovane, teda dataset s daným názvom a od danej organizácie je považovaný aktuálne za jeden dataset, bez ohľadu na zdrojové dáta, obdobia, formát datasetu. Napriek tomu napr. dataset *Vestník verejného obstarávania*⁴³ od Úradu pre verejné obstarávanie je vždy nanovo zverejňovaný mesačne. Bolo by vhodné jednoznačne zdefinovať, aké dáta sú považované za jeden dataset – návrhom cez Komisiu pre štandardizáciu informačných systémov verejnej správy (ÚPVII).
4. Na portáli data.gov.sk sú zverejnené aj datasety bez licencie alebo s nevhodnou, obmedzujúcou licenciou. V takých prípadoch nejde o otvorené dáta. Bolo by užitočné, keby portál upozornil povinné osoby pri zverejnení datasetu na chýbajúcu alebo nevhodnú licenciu a naviedol povinné osoby vybrať si zo zoznamu vhodných licencií. Ku zverejneným datasetom bez licencie treba doplniť licenciu.
5. Niektoré formáty datasetov, zobrazené na portáli data.gov.sk v záložke Datasets, filter Formáty, nie sú formátmi skutočných datasetov, ale pomocnými súbormi k datasetom (napr. XSD ku XML). Také pomocné súbory by mohli byť špeciálne označené, aby bolo jasné, ku ktorému formátu datasetov patria. Niektoré formáty sú nesprávne alebo duplicitne pomenované alebo ide kombináciu formátov (napr. XML, xml, xml.net, xslx, .xlsx, PDF, pdf, HTML, html, htm, CSV, csv, csw). Tieto formáty treba zjednotiť. Povinná osoba by mala mať pri zverejnení datasetu zoznam formátov na výber a pri zadaní nového formátu by mala prebehnúť kontrola, či už taký formát nie je zadaný podobnou skratkou. Medzi formátmi môžu byť chybne uvedené prepojenia (api), alebo hypertextové odkazy na iné sídlo mimo data.gov.sk (url), tie by mohli byť odlišené od súborov na data.gov.sk.
6. Existuje zrozumiteľné hodnotenie otvorených dát hviezdčkami⁴⁴, známe na Slovensku aj vo svete. Kvalita poskytovaného datasetu je v tomto význame zadaná vo vyhláske ÚPVII

41 https://www.slovensko.sk/sk/agendy/agenda/_ministerstva-a-ostatne-ustredne1

42 <https://www.slovensko.sk/sk/institucie>

43 Napr. <https://data.gov.sk/dataset/vestnik-verejneho-obstaravania-20205>

44 <https://5stardata.info/cs/>

č. 78/2020 Z. z. (Vyhláška o štandardoch pre informačné technológie verejnej správy)⁴⁵, § 38 až § 40. Na zverejňovanie otvorených dát sa vzťahuje tiež Národná koncepcia informatizácie verejnej správy⁴⁶, Stratégia a akčný plán sprístupnenia a používania otvorených údajov verejnej správy⁴⁷. Minimálna požadovaná úroveň pre otvorené dáta je úroveň 3 hviezdičiek. Portál data.gov.sk by mohol automaticky hodnotiť zverejnené datasety hviezdičkami na základe platného predpisu, inými slovami: podľa typu licencie, formátu datasetu, priradeného referencovateľného identifikátora a existencie prepojenia na iný dataset cez referencovateľný identifikátor. Toto hodnotenie by sa mohlo prehľadne sumarizovať v záložke Štatistiky.

7. Zaviesť striktné požiadavky na zverejnené datasety – portál data.gov.sk by mohol umožniť zverejnenie len tých datasetov, ktoré spĺnia aspoň 3 hviezdičky – čo by bolo v súlade s platnými predpismi (vymenované v predošlom bode).
8. Zaviesť sprístupnenie hodnotenia a komentovania datasetov aj pre neprihlásených používateľov portálu data.gov.sk, najlepšie cez automatizované hlásenia. Dôvod: jednoduchosť, prehľadnosť pre používateľov, spätná väzba pre poskytovateľov datasetov.
9. Verejnosť a viacerí dátoví kurátori organizácií by chceli vidieť štatistiku prístupov na jednotlivé popisné HTML stránky datasetov na portáli data.gov.sk. Tým by sa tiež zhodnotil dopyt verejnosti po najžiadanejších datasetoch. Avšak vytvoriť objektívnu štatistiku prístupov je ťažké. Portál by mohol pri datasetoch zobrazovať jednoduchú grafickú štatistiku prístupov podľa času.
10. Portál data.gov.sk by mohol zobrazovať štatistiku využitia jednotlivých podaní⁴⁸ verejnosťou.
11. Pri niektorých zobrazeniach na portáli data.gov.sk sú viditeľné nedostatky vo webdizajne portálu. Napr. slabý farebný kontrast v HTML/CSS – nezadaná farba písma spolu s farbou pozadia v tabuľke *Dodatočné informácie* pri kliknutí na hociktorý dataset⁴⁹. Také nedostatky treba odstrániť, aby portál data.gov.sk bol v súlade s odporúčaniami WCAG 2.0. Pri úprave webdizajnu sa treba riadiť Jednotným dizajnom manuálom elektronických služieb⁵⁰.
12. Portál data.gov.sk by mohol byť doplnený o ďalšie štatistiky alebo analytické prehľady (všetky podľa tejto analýzy). Ďalšia analýza by mohla byť automatizovaná, objektívnejšia, bez rozdielov v metodike medzi rokmi.
13. Uľahčiť pridávanie aplikácií postavených na otvorených dátach. Pridávať nové aplikácie iniciatívne podľa prieskumu dopytu verejnosti po najžiadanejších datasetoch a API, s aktívnou pomocou dátových kurátorov.

45 <https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2020/78/20200501>

46 <https://www.vicempremier.gov.sk/sekcie/narodna-koncepcia-informatizacie-verejnej-spravy-nikvs/index.html>

47 https://www.slovensko.sk/_img/CMS4/strategia_otvorene_udaje.pdf

48 <https://data.gov.sk/published-services>

49 napr. <https://data.gov.sk/dataset/pocet-rozhodnuti-podani-a-notifikacii-zaslanych-cez-upvs>

50 <https://idsk.gov.sk>

Odporúčania pre dátovú kanceláriu ÚPVII

1. Po vzore ČR⁵¹ monitorovať „časté zlé praktiky“ a zverejňovať krátke návody, popisujúce konkrétny nedostatok a kroky potrebné pre nápravu.
2. Analyzovať dôvody medziročného poklesu počtu datasetov (podľa kap. 1.1) a vykonať podporné kroky, aby v budúcnosti k poklesu počtov nedochádzalo. Zverejnenie datasetov by malo byť v zmysle predpisov trvalé (zverejniť a archivovať).

Odporúčanie pre povinné osoby – poskytovateľov datasetov

*Metodická príručka pre povinné osoby – ako zverejňovať otvorené dáta (2015)*⁵² obsahuje všetko potrebné pre správne zverejňovanie datasetov. Každá povinná osoba, u ktorej sa vyskytol nejaký nedostatok opísaný v tejto analýze, by mala postupovať podľa metodической príručky a napraviť nedostatky.

Poznámka

Úrad podpredsedu vlády SR pre investície a informatizáciu (ÚPVII) sa mení na Ministerstvo investícií, regionálneho rozvoja a informatizácie SR (MIRRI SR) k 1.10.2020.

51 <https://opendata.gov.cz/špatná-praxe:start>

52 https://www.slovensko.sk/img/CMS4/Navody/Metodicka_prirucka_pre_povinne_osoby.pdf