

Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát 2019

na základe

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

Bratislava, 28.6.2019



Obsah

Úvod.....	3
1 Kvantitatívna analýza	4
1.1 Počet zverejnených datasetov.....	5
1.2 Tagy (kľúčové slová).....	6
1.3 Licencie datasetov	7
1.4 Formáty datasetov	7
1.5 Aktualizácia datasetu.....	8
1.6 Aktuálnosť datasetu.....	9
1.7 Veľkosť datasetu.....	10
1.8 Používatelia datasetov	10
2 Kvalitatívna analýza	11
2.1 Hodnotenie datasetov	11
2.2 Global open data index.....	14
2.3 Open data barometer	14
2.4 Open data maturity index	15
2.5 Open data watch	15
2.6 Nástroje	15
2.7 Aplikácie	15
2.8 Užitočnosť datasetu.....	16
Záver	18

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 2017 – 2019*, schválený Uznesením vlády SR č. 104/2017, v kapitole *1.1 Portál otvorených dát – data.gov.sk* obsahuje úlohu 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. Zodpovedný: splnomocnenec vlády SR pre rozvoj občianskej spoločnosti. Termín: každoročne do 30. júna.

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. Špeciálnym datasetom je zoznam zverejnených datasetov.

Predchádzajúce analýzy zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát boli realizované za roky 2017, 2018. 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 neskúma zverejnené datasety verejnej správy mimo ústredných orgánov štátnej správy, ani zverejnené datasety územnej samosprávy.



Obrázok 1: <https://data.gov.sk>

Správcom portálu <https://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 data.gov.sk poskytuje okrem zverejnených datasetov aj vyhľadávanie a filtrovanie datasetov podľa rôznych kritérií, návody a iné súvisiace informácie.

1 Kvantitatívna analýza

V tejto kapitole **sumarizujeme počty datasetov podľa viacerých kritérií** a porovnávame podľa rokov zverejnenia.

Je možné zistiť rozdielne **absolútne počty datasetov** pre dané prípady **v závislosti od vyhľadávacieho 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. Kvôli zjednodušeniu považujeme dataset s daným názvom a od danej organizácie za jeden, bez ohľadu na zdrojové dáta, obdobia, formát datasetu. Napriek tomu napr. dataset *Vestník verejného obstarávania*¹ od Úradu pre verejné obstarávanie je vždy nanovo zverejňovaný mesačne, aj keď ide o rovnaký dataset. Aj metodika predchádzajúcich analýz sa môže líšiť, preto absolútne počty datasetov sú skôr orientačné. Dáta v tabuľke 1 boli zisťované k dátumom: 27.08.2013, 08.03.2015, 31.03.2017, 30.06.2018, 25.06.2019². Portál data.gov.sk v záložke Štatistiky uvádza sumárny počet datasetov podľa organizácií³.

Táto analýza neskúma zverejnené datasety verejnej správy mimo ústredných orgánov štátnej správy, ani zverejnené datasety územnej samosprávy. V tabuľke uvádzame zoznam všetkých ústredných orgánov štátnej správy a kvôli úplnosti pridávame sumárne počty za ostatné organizácie verejnej správy a organizácie územnej samosprávy. V absolútnych počtoch sa môže vyskytnúť čiastková chyba, ak daná podriadená organizácia bola v jednom roku pripočítaná k príslušnému ústrednému orgánu štátnej správy (ministerstvu) a v inom roku bola pripočítaná k ostatným organizáciám verejnej správy. Súčet všetkých datasetov v roku sa však týmto nezmení.

Portál data.gov.sk v záložke Datasety zobrazuje sumárny **počet datasetov podľa tagu (kľúčového slova)** pre ľahšie vyhľadávanie. 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.

Portál data.gov.sk v záložke Štatistiky poskytuje sumárny **počet datasetov podľa typu licencie**. Tieto počty po prepočte na percentá uvádzame vo forme grafu. Celkový počet datasetov zistený pri licenciách sa však nezohoduje s celkovým počtom licencií zisteným podľa organizácií. Napriek tomu, že portál data.gov.sk plní funkciu portálu otvorených dát, sú zverejnené aj datasety bez uvedenia licencie alebo s obmedzujúcou licenciou.

Portál data.gov.sk v záložke Datasety zobrazuje sumárny **počet datasetov podľa formátu súboru**. Tieto počty v prepočte na percentá uvádzame v grafe. Celkový počet datasetov zistený pri formátoch nesedí s celkovým počtom datasetov zisteným pri organizáciách, pretože niektoré datasety sú zverejnené vo viacerých formátoch alebo pre rôzne zdrojové dáta. Niektoré formáty uvedené na portáli nie sú formátmi skutočných datasetov, ale pomocnými súbormi k datasetom (napr. XSD ku

¹napr. <https://data.gov.sk/dataset/vestnik-verejneho-obstaravania-20195>

²<https://data.gov.sk/organization>

³Sumárne počty v štatistike data.gov.sk (na obrázku) sa nezohodujú so skutočným počtom datasetov po kliknutí.

XML). Niektoré formáty sú nesprávne alebo duplicitne pomenované. Medzi formáty sú zaradené aj hypertextové odkazy na iné sídlo mimo data.gov.sk.

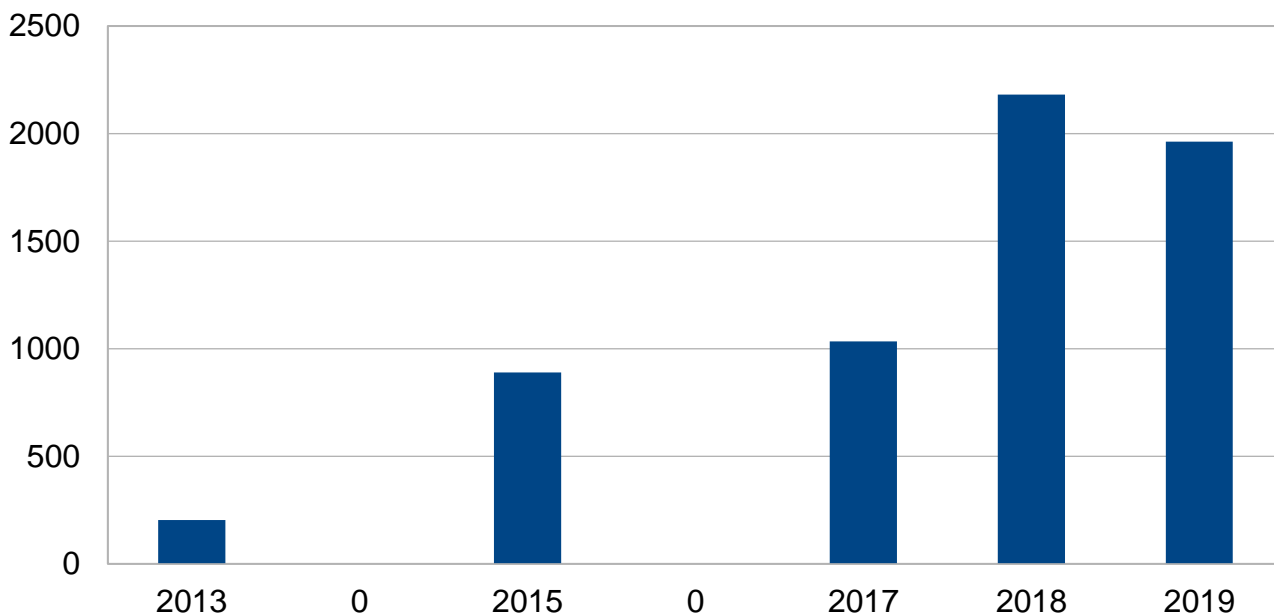
1.1 Počet zverejnených datasetov

Tabuľka 1: Počet zverejnených datasetov podľa organizácií a rokov

Ministerstvo hospodárstva SR	7	8	9	17	17
Ministerstvo kultúry SR	10	11	11	35	25
Ministerstvo obrany SR	3	4	4	19	8
Ministerstvo pôdohospodárstva a rozvoja vidieka SR	65	68	67	209	457
Ministerstvo práce, sociálnych vecí a rodiny SR	5	5	5	5	13
Ministerstvo spravodlivosti SR	5	8	8	17	16
Ministerstvo školstva, vedy, výskumu a športu SR	5	24	24	34	26
Ministerstvo vnútra SR	24	40	41	253	61
Ministerstvo zahraničných vecí a európskych záležitostí SR	14	14	16	25	25
Ministerstvo zdravotníctva SR	9	18	95	184	262
Ministerstvo životného prostredia SR	9	9	22	41	25
Národný bezpečnostný úrad				4	23
Protimonopolný úrad SR				7	7
Správa štátnych hmotných rezerv SR				5	7
Štatistický úrad SR	9	586	606	1 101	665
Úrad geodézie, kartografie a katastra SR	2	2	2	2	2
Úrad jadrového dozoru SR				6	7
Úrad podpredsedu vlády SR pre investície a informatizáciu	0	23	12	36	39
Úrad pre normalizáciu, metrológiu a skúšobníctvo SR				3	7
Úrad pre verejné obstarávanie	7	7	46	16	79
Úrad priemyselného vlastníctva SR				19	8
Úrad vlády SR	7	11	11	13	15
Generálna prokuratúra SR				8	8
Kancelária Národnej rady SR					0
Najvyšší kontrolný úrad SR					2
Národná banka Slovenska					
organizácie územnej samosprávy spolu	0	18	22	67	68
ostatné organizácie verejnej správy spolu	4	1	1	3	11
Spolu	204	889	1 034	2 181	1 962

Dáta z roku 2019 sú k 25.06.2019. Celkový počet organizácií v štatistike portálu data.gov.sk je 83.

V absolútnom počte zverejnených datasetov vedie Štatistický úrad SR, čo je pochopiteľné. Z územnej samosprávy dlhodobo vedie mesto Prešov.

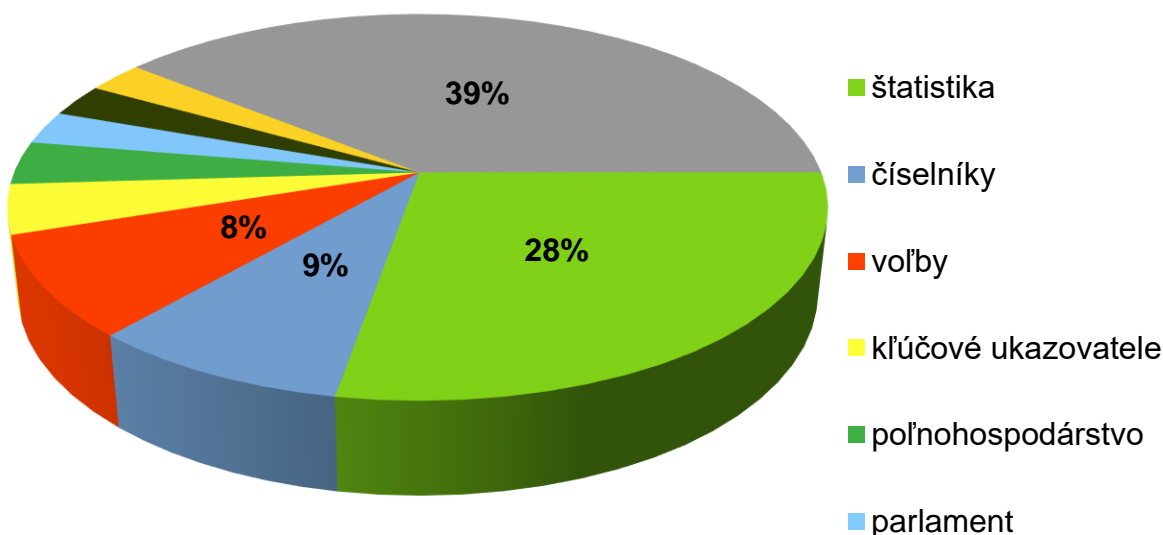


Graf 1: Celkový počet datasetov podľa rokov

Poznámka: Metodika určenia počtu datasetov v jednotlivých rokoch môže byť rozdielna.

1.2 Tagy (kľúčové slová)

Portál data.gov.sk umožňuje vyhľadávať datasety podľa tagu⁴ (kľúčového slova), čo je veľmi užitočná funkcia. Portál ponúka niekoľko obľúbených, často používaných tagov.

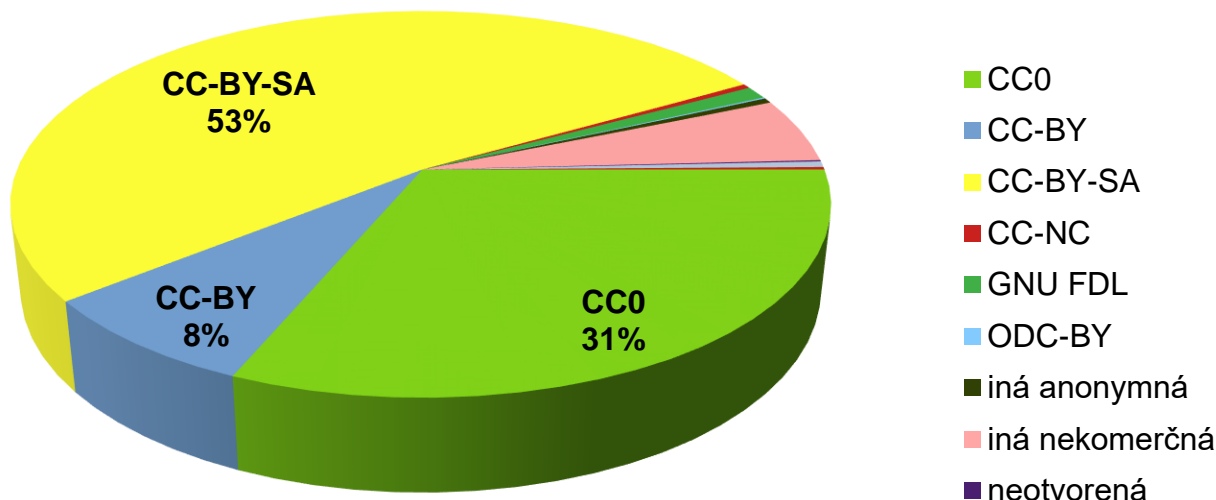


Graf 2: Počet datasetov podľa tagov (kľúčových slov)

Najväčší podiel majú datasety s tagom *štatistika*, čo je pochopiteľné lebo práve Štatistický úrad SR má najväčší počet zverejnených datasetov. Záujem používateľov o datasety volieb zrejme bude časovo variabilný podľa dátumu volieb.

⁴https://data.gov.sk/dataset?_tags_limit=0

1.3 Licencie datasetov

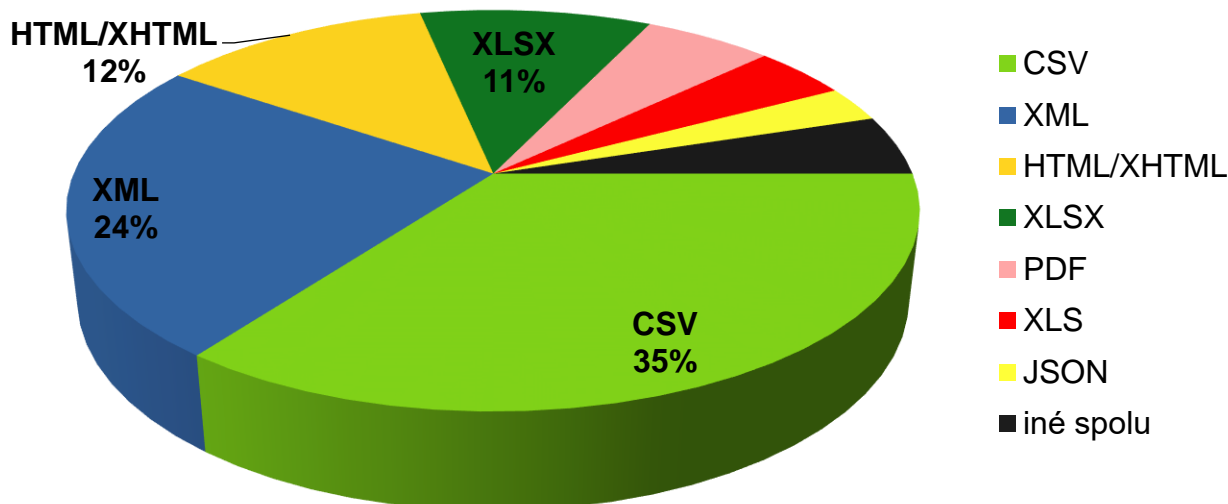


Graf 3: Typy licencií

V roku 2017 najväčší podiel pri zverejnených datasetoch mala licencia CC-BY-SA⁵. V roku 2018 výrazne vzrástol počet datasetov s licenciou CC0. V roku 2019 je opäť na prvom mieste licencia CC-BY-SA a licencia CC0 sa drží na druhom mieste.

Poskytovateľ datasetov, ktorý neuvádza licenciou, je Úrad vlády SR. Viacerí poskytovatelia malého počtu datasetov uvádzajú obmedzujúce licencie.

1.4 Formáty datasetov



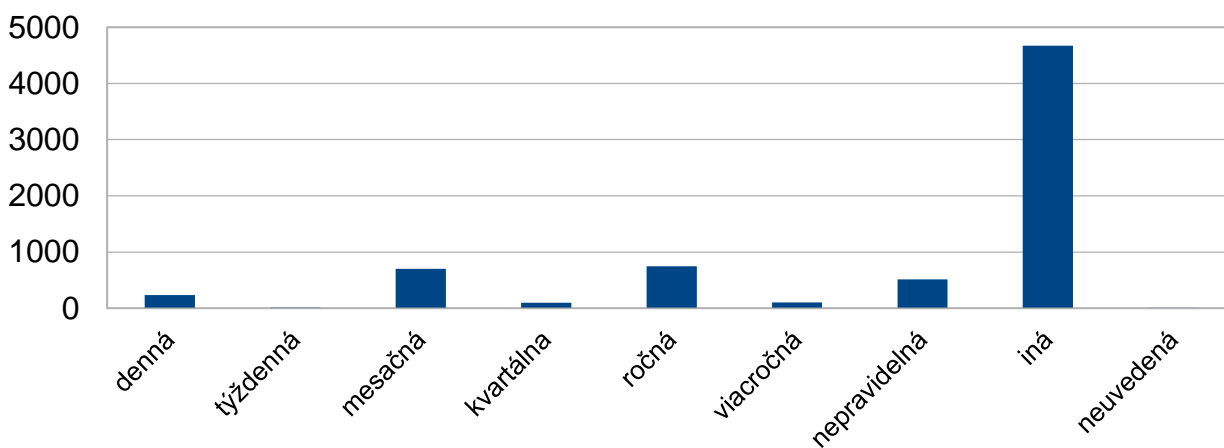
Graf 4: Formáty datasetov

V roku 2018 dominoval formát XML. V roku 2019 je najpoužívanejším formát CSV. Formát PDF má podiel 6 %, čo je stále veľa vzhľadom na horšiu strojovú spracovateľnosť tohto formátu. Formát

⁵Prehľad licencií Creative Commons uvádza stránka https://sk.creativecommons.org/?page_id=6

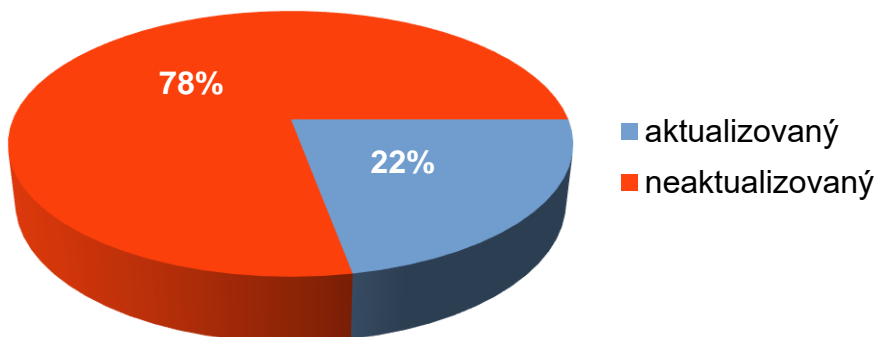
XLSX by mal byť zverejňovaný spolu s alternatívnym formátom ODS, ten má ale podstatne nižší podiel ako XLSX. Relatívne malý počet datasetov je sprístupnený v niektorom z viacerých zastaraných alebo nevhodných formátov, napr. XLS, RTF, DOCX, DOC, RAR, JPG.

1.5 Aktualizácia datasetu



Graf 5: Počet datasetov podľa pravidelnosti aktualizácie

Vychádzame zo zoznamu datasetov od prevádzkovateľa (NASES)⁶, absolútne počty datasetov sa líšia od predchádzajúcich kapitol. V grafoch 5, 6, 7, 8, 10 započítaváme každý jednotlivý dataset, aj keď je zverejnený viacnásobne z dôvodu rozdielnych zdrojových dát, období alebo viacerých formátov. Portál data.gov.sk neposkytuje všetky štatistiky potrebné pre túto analýzu, pričom portál ako zdroj štatistík uprednostňujeme pre jednoduchosť, prístupnosť. Jednotná metodika len podľa zoznamu od prevádzkovateľa by znamenala o dosť náročnejšie spracovanie dát. V roku 2018 pri väčšine datasetov bola *neuvedená* pravidelnosť aktualizácie. V roku 2019 pri väčšine datasetov je *iná* pravidelnosť aktualizácie (iná perióda, možno aj žiadna aktualizácia).



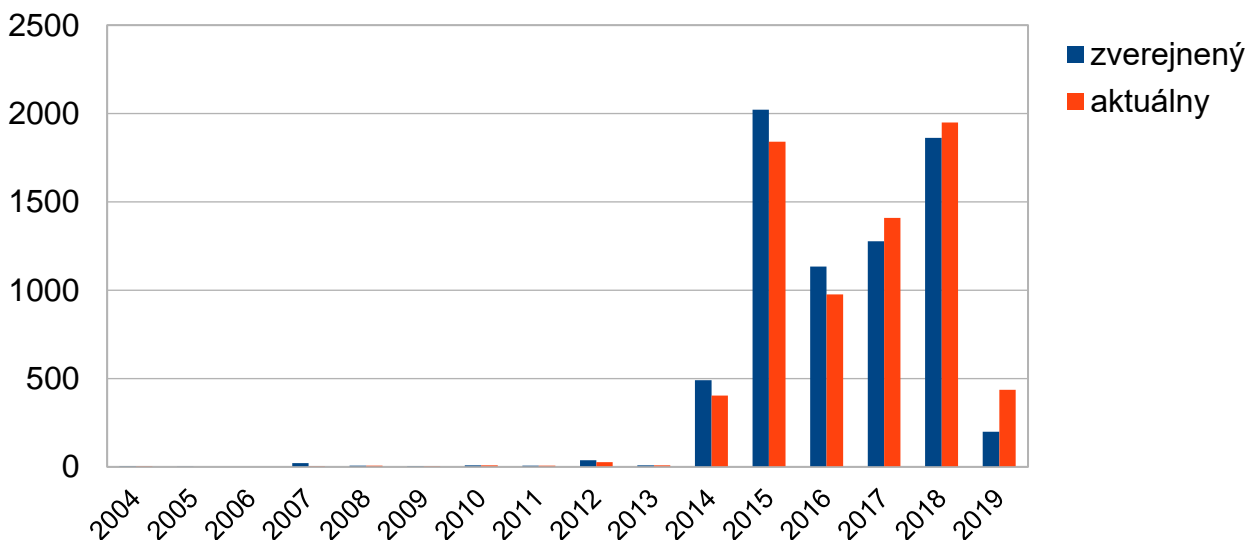
Graf 6: Podiel datasetov s aktualizáciou

Podiel aktualizovaných datasetov je príliš nízky, podobne to bolo aj v roku 2018. Možné príčiny: ešte neuplynula lehota platnosti datasetu do aktualizácie (*iná*); dataset nebol aktualizovaný,

⁶Súbor: modDbExport_activePublic.xlsx (04.04.2019)

aj keď mal byť; bol zverejnený nový dataset namiesto aktualizácie; dataset je zverejnený inde, na portáli je len URL adresa datasetu a portál nemá informáciu o aktualizácii.

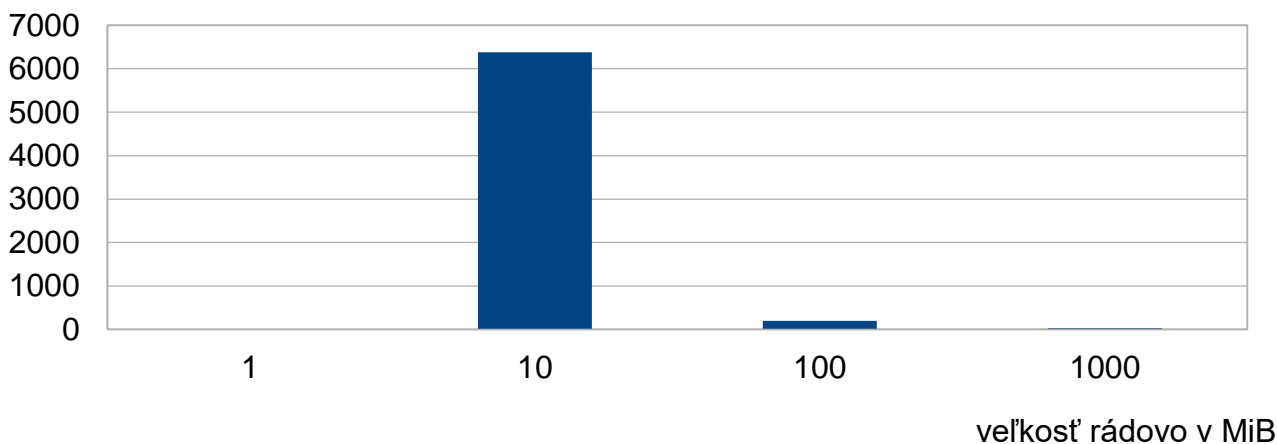
1.6 Aktuálnosť datasetu



Graf 7: Počet aktuálnych datasetov podľa rokov

Z grafu vyplýva, že počet aktuálnych datasetov je od roku 2017 vyšší ako pôvodne zverejnených datasetov v danom roku. Nepriamo možno predpokladať, že v priemere ak je dataset aktualizovaný, je aktualizovaný do 3 rokov. Niektoré datasety môžu byť jednorázové, bez potreby aktualizácie. Červený stĺpec je maximum z roku zverejnenia a z roku aktualizácie. V zozname datasetov sú niektoré s dátumom zverejnenia pred rokom 2011, kedy sa začína s Iniciatívou pre otvorené vládnutie: registre odborne spôsobilých osôb (regionálne úrady verejného zdravotníctva).

1.7 Veľkosť datasetu



Graf 8: Počet datasetov podľa veľkosti

Najväčší zverejnený dataset je *Register Adries – Register vchodov (orientačných čísiel)* od Ministerstva vnútra SR, s veľkosťou 938 MiB. Časť datasetov nemá uvedenú veľkosť. Veľkosť všetkých datasetov spolu je 23 GiB.

1.8 Používatelia datasetov

Na portáli data.gov.sk v záložke Štatistiky sa uvádza počet registrovaných používateľov portálu (používateľov datasetov) spolu 3 811. Keďže portál s datasetmi je verejne prístupný, používateľ nemusí byť registrovaný. Všetkých používateľov bude viac. Registrovaný používateľ dostáva navyše napríklad notifikácie o datasetoch (nových, aktualizovaných, zmazaných a podobe), metadáta k datasetom.

Portál data.gov.sk sprístupňuje služby priamo, alebo cez *Ústredný portál verejnej správy* (<https://www.slovensko.sk>), pomocou Katalógu služieb. Portál data.gov.sk informuje o týchto poskytovaných službách⁷:

- ⑩ poskytovanie dát v otvorených formátoch
- ⑩ poskytovanie metadát k poskytovaným dátam alebo dátovému zdroju
- ⑩ grafická vizualizácia poskytovaných dát
- ⑩ geografická prezentácia poskytovaných dát
- ⑩ vyhľadávanie dát podľa tagov (kľúčových slov)
- ⑩ vyhľadávanie dát podľa organizácie (poskytovateľa)
- ⑩ vyhľadávanie dát podľa geografickej lokácie

Verejnoscť môže využiť nasledovné podania:

- ⑩ žiadosť o sprístupnenie informácií na účely opakovaného použitia podľa zákona o slobodnom prístupe k informáciám

⁷<https://data.gov.sk/published-services>

- ⑩ návrh na zverejnenie údajov na Portáli otvorených dát
- ⑩ žiadosť o registráciu aplikácie na Portáli otvorených dát
- ⑩ podnet na úpravu údajov zverejnených na Portáli otvorených dát
- ⑩ žiadosť o zvýšenie kapacity na uloženie datasetov na Portáli otvorených dát (neposkytuje sa)
- ⑩ žiadosť o pridelenie zvýhodneného prístupu k údajom zverejneným na Portáli otvorených dát (neposkytuje sa)

2 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.

2.1 Hodnotenie datasetov

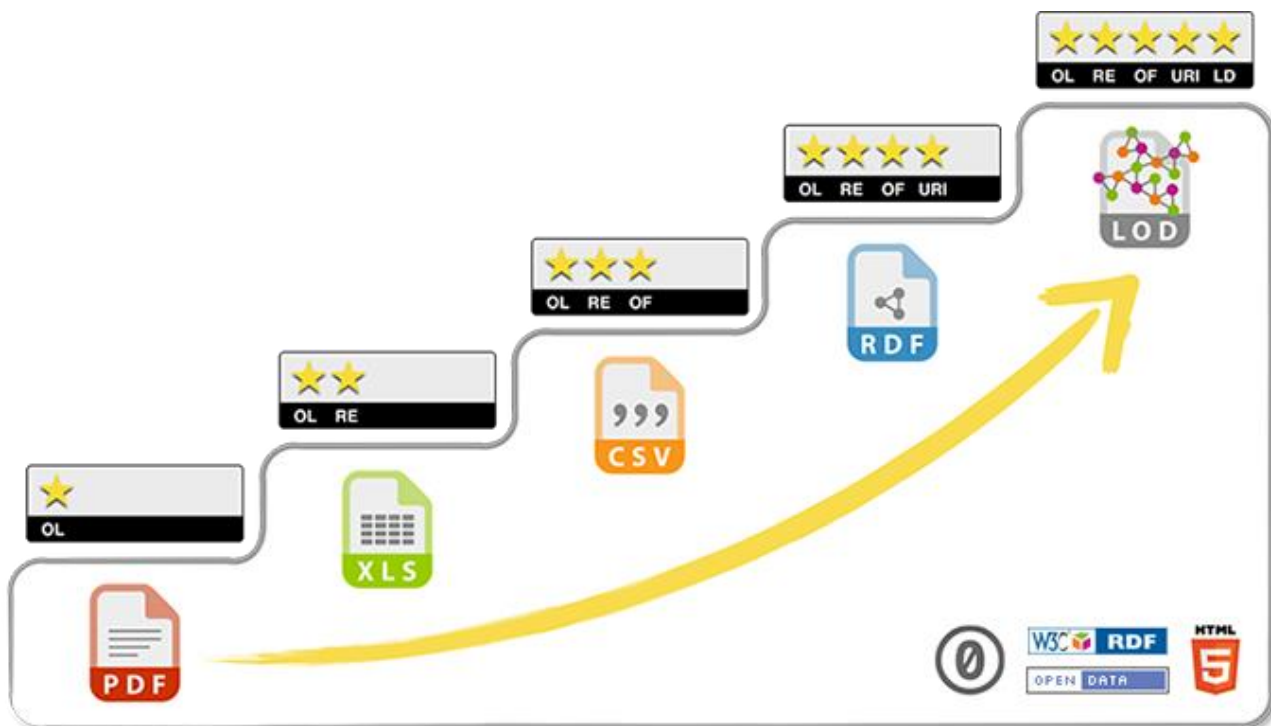
Tim Berners-Lee, vynálezca webu a iniciátor prepojených dát, navrhol hodnotenie otvorených dát hviezdikami⁸:

- * 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 výnose MF SR č. 55/2014 Z. z. (Výnos o štandardoch pre informačné systémy verejnej správy)⁹, § 51 až § 53.

⁸<https://5stardata.info/cs/>

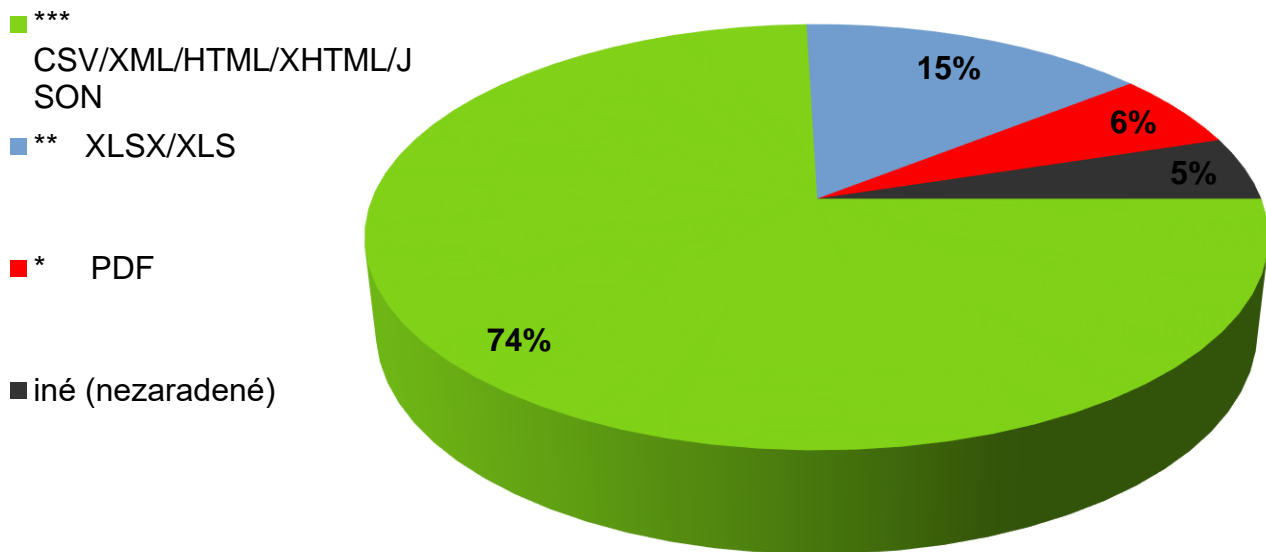
⁹<https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2014/55/20190601>



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.

Najprv ohodnotíme formáty zverejnených datasetov hviezdikami, podľa dát pre graf 4 (formáty datasetov). Ľahko sa dajú určiť hodnotenia jednou, dvomi, tromi hviezdikami.



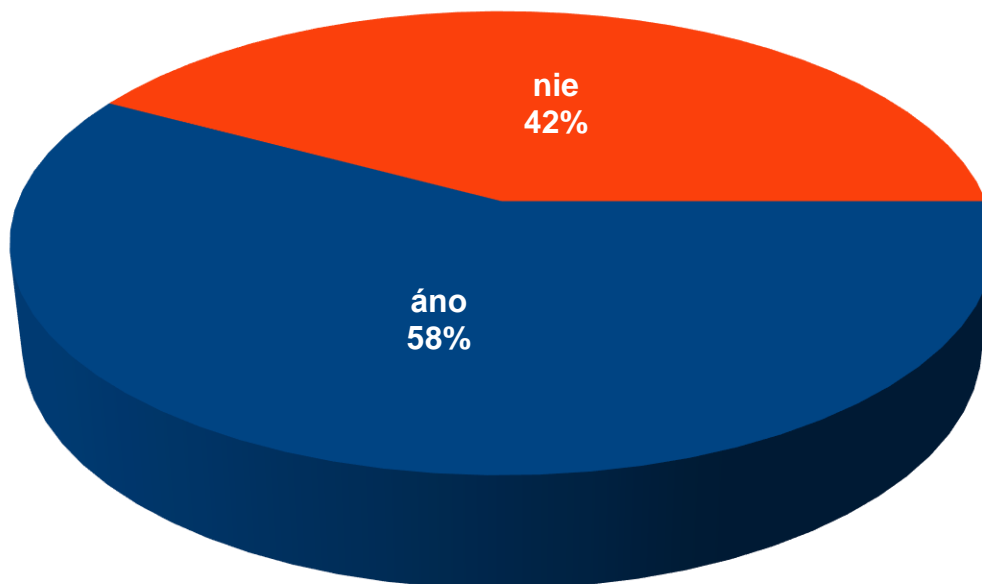
Graf 9: Hodnotenie formátu hviezdikami

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

Teraz urobme prienik vhodnej licencie datasetu a vhodného formátu datasetu, na základe zoznamu datasetov od prevádzkovateľa (NASSES)¹⁰. Minimálna požadovaná úroveň pre otvorené dáta je úroveň 3 hviezdíček.

¹⁰Súbor: modDbExport_activePublic.xlsx (04.04.2019)

Za



Graf 10: Podiel datasetov s aspoň 3 hviezdčkami podľa licencie a formátu

vhodné licencie považujeme: CC0, CC-BY, CC-BY-SA, ODC-BY, GNU FDL, iné otvorené. Za vhodné formáty považujeme: CSV, XML, HTML, XHTML, JSON, RDF.

Úroveň 4 hviezdčiek spĺňajú tie datasety, ktoré majú 3 hviezdčičky a obsahujú aj metadáta a referencovateľný identifikátor (permanentná URI/URL adresa). Predpokladáme, že datasety na portáli data.gov.sk sú určené na trvalé zverejnenie, pričom portál poskytuje ku každému datasetu permanentnú URI/URL adresu, takže možno nepriamo usúdiť, že datasety z grafu 10, ktoré majú 3 hviezdčičky (58 %), majú potom 4 hviezdčičky. Úroveň 5 hviezdčiek spĺňajú tie datasety, ktoré majú 4 hviezdčičky a obsahujú v dátach prepojenia referencovateľnými identifikátormi na iné datasety. Toto nevieme dostatočne rýchlo vyhodnotiť z dostupných informácií o datasetoch.

Séria 6 datasetov od Ministerstva životného prostredia SR alebo jeho podriadených organizácií je dobrým príkladom 5-hviezdčkových datasetov. Sú zverejnené s licenciou CC-BY, vo formáte RDF a sú prepojené, čo je navyše indikované tagom *Linked Data*.

2.2 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.

2.3 Open data barometer

<https://opendatabarometer.org>

Slovensko nie je zaradené do rebríčka v poslednom sledovanom období 2017. V predchádzajúcom období 2016 sa nezobrazujú dáta. V období 2015 sa Slovensko umiestnilo na 36. mieste s 37,16 bodmi zo 100 v rebríčku 92 štátov (Spojené kráľovstvo na 1. mieste so 100 bodmi).

2.4 Open data maturity index

<https://www.europeandataportal.eu>

Slovensko sa umiestnilo v poslednom sledovanom období 2018¹¹ na 9. mieste so 74 % v rebríčku 31 sledovaných európskych štátov (Írsko na 1. mieste s 88 %). Metodika je založená na posúdení 4 dimenzií otvorených dát: politika, portál, kvalita, dopad. Aj v roku 2015 bolo Slovensko na 9. mieste.

2.5 Open data watch

<https://opendatawatch.com>

Slovensko sa umiestnilo v poslednom sledovanom období 2018¹² 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 2018¹³ je na základe dát Štatistického úradu SR¹⁴.

2.6 Nástroje

Na portáli data.gov.sk v záložke Nástroje je dostupný iba 1 nástroj:

SPARQL editor

Tento nástroj slúži na vytváranie dopytov v jazyku SPARQL (*SPARQL protocol and RDF query language*) a prezentáciu získaných dát priamo vo webovom prehliadači. SPARQL slúži na prácu s RDF (*resource description framework*) dátami, čo sú metadáta ku zverejneným datasetom.

2.7 Aplikácie

Na portáli data.gov.sk v záložke Aplikácie sú dostupné 4 aplikácie (1 nová z roku 2019):

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

¹¹<https://www.europeandataportal.eu/en/highlights/open-data-maturity-report-2018>

¹²<http://odin.opendatawatch.com/Report/rankings?sortOrder=&appConfigId=5>

¹³<http://odin.opendatawatch.com/Report/countryProfile/SVK?appConfigId=5>

¹⁴<http://www.statistics.sk>

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>

Revízia základných sídelných jednotiek

Aplikácia pre podporu prípravy podkladov pre sčítanie obyvateľstva, domov a bytov.

Slovenská agentúra životného prostredia, 2016, <http://rzs.j.enviroportal.sk/mapov-zobrazenie>

2.8 Užitočnosť datasetu

Prieskum dopytu verejnosti po najžiadanejších datasetoch sa vykonáva každoročne. Výsledky prieskumu sa však stali dlhodobým prehľadom najžiadanejších datasetov¹⁵, pretože v *Strednodobej hodnotiacej správe Akčného plánu Iniciatívy pre otvorené vládnutie v Slovenskej republike na roky 2017 – 2019*¹⁶ (str. 19) sa píše, že množstvo datasetov, ktoré boli požadované v roku 2017, nie je zverejnených k 31.07.2018. Zo žiadaných datasetov, ktoré zatiaľ neboli zverejnené vyberáme napr.:

- Cestovné poriadky verejnej dopravy, aktuálne polohy vozidiel a aktuálne informácie o verejnej doprave, štatistika vyťažnosti, zoznamy zastávok a staníc: MHD, autobusy, vlaky (Ministerstvo dopravy a výstavby SR)
- Legislatíva v strojovo spracovateľnom formáte, právne predpisy a návrhy právnych predpisov v strojovo spracovateľnom formáte (Ministerstvo spravodlivosti SR)
- Rozpočty štátnych podnikov, verejnej správy, územnej samosprávy (Ministerstvo financií SR)
- Živnostenský register (Ministerstvo vnútra)
- Zoznam štátnych budov, otváracie hodiny, telefónne čísla (Úrad vlády SR)
- Katastrálne mapy, kataster, listy vlastníctva (Úrad geodézie, kartografie a katastra SR)
- Zoznam vrcholov, zoznam minerálnych prameňov, zoznam riek, dáta o znečistení životného prostredia, SHMÚ dáta (Ministerstvo životného prostredia SR)

Zverejnenie niektorých datasetov je aspoň naplánované na najbližšie obdobie.

Pri najžiadanejších datasetoch vyžiadaných verejnosťou v prieskume dopytu verejnosti po najžiadanejších datasetoch predpokladáme, že užitočnosť takého datasetu je objektívne vysoká. Zverejnené sú aj datasety (konkrétne dáta v datasetoch) s nízkou užitočnosťou, zrejme preto, že povinná osoba (organizácia) má povinnosť zverejniť taký dataset, ale nie je s tým stotožnená (jej zodpovedný pracovník) a kontrola je len formálna, či bol dataset zverejnený. Jeden príklad: *Zoznam využívaného otvoreného softvéru v gescii MŠVVaŠ SR*¹⁷. Dataset je 2-hviezdičkový (licencia CC0, formát XLSX). Je aktualizovaný 28.7.2017 (viac ako 1 rok starý), pričom dáta v tomto datasete podliehajú zmenám na mesačnej báze. Pravidelnosť aktualizácie: 1 rok (po termíne). Dáta by mohli byť užitočné, keby ku vymenovanému softvéru bol pridaný počet inštalácií, účel použitia, miesto

¹⁵<https://docs.google.com/spreadsheets/d/13uVKtcS4vDHZISeLZvv-0E3Cefmutc8P-CVk0cEf1RM/edit#gid=0>

¹⁶https://www.minv.sk/swift_data/source/rozvoj_obcianskej_spolocnosti/otvorene_vladnutie/2018/strednodoba_hodnoti_aca_sprava/2018_07_31-OGP%202017-2019_strednodoba%20sprava_verzia%203_LB.pdf

¹⁷<https://data.gov.sk/dataset/zoznam-vyuzivaneho-otvoreneho-softveru-v-gescii-msvvas-sr>

použitia (sumárne) a podobne. Dataset by mohol byť napr. podkladom pre analýzu hodnoty za peniaze, podľa Národnej koncepcie informatizácie verejnej správy (NKIVS).

Záver

Na záver uvádzame **odporúčania na najbližšie obdobie:**

- **Prieskum dopytu verejnosti** – pokračovať v prieskume dopytu verejnosti po najžiadanejších datasetoch otvorených dát.
- **Monitorovať a vyhodnocovať zverejnenie dopytu verejnosti** po najžiadanejších datasetoch orgánov územnej samosprávy.
- **Zvýšiť kvalitu datasetov** – portál data.gov.sk sa už niekoľko rokov postupne plní datasetmi. V nasledujúcich rokoch by sme sa mali viac zaujímať o kvalitu datasetov.
- **Rozšírenie záberu dát** – vzhľadom na hodnotenie otvorených dát hviezdikami (kap. 2.1), na pripravovanú *Stratégiu digitálnej transformácie Slovenska 2030*¹⁸ a na pripravovaný program *Digitálna Európa 2021 – 2027*¹⁹ je vhodné, aby sa pozornosť venovala týmto témam: datasety samosprávy (obce, mestá, kraje), otvorené prepojené dáta (interoperabilita), otvorené veľké dáta, otvorené dáta a GDPR, súkromie, etika.
- **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, inteligentné mesto a internet vecí, otvorený softvér a otvorený hardvér.
- V novom akčnom pláne by mali byť témy o otvorených dátach previazané aj na kapitolu **Otvorené vzdelávanie a otvorená veda**.

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. Pridávame odporúčania, ako ich vyriešiť:

1. Na portáli data.gov.sk v záložke Štatistika na obrázku sa sumárne počty datasetov nezhodujú so skutočným počtom datasetov po kliknutí na daný riadok v obrázku. Zrejme číslo na obrázku nie je aktuálne. Mal by sa zobrazovať všade aktuálny stav datasetov.
2. Je užitočné, že portál data.gov.sk zverejňuje datasety od každej organizácie verejnej správy a územnej samosprávy. Avšak bolo by prehľadnejšie, ak by portál data.gov.sk umožňoval zobrazovať zoznam ústredných orgánov štátnej správy (ministerstvá a ústredné úrady) a k nim priradené podriadené organizácie. Napríklad Regionálne úrady verejného zdravotníctva

¹⁸<https://www.slov-lex.sk/legislativne-procesy/SK/LP/2019/205>

¹⁹EC(2018) 434 final, <https://www.slov-lex.sk/legislativne-procesy/SK/LP/2018/687>

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 za jeden dataset, bez ohľadu na zdrojové dáta, obdobia, formát datasetu. Napriek tomu napr. dataset *Vestník verejného obstarávania*²⁰ od Úradu pre verejné obstarávanie je vždy nanovo zverejňovaný mesačne, aj keď ide o rovnaký dataset. Bolo by vhodné jednotne nastaviť a dodržiavať pravidlo, aké dáta sú považované za jeden dataset.
4. Na portáli data.gov.sk sú zverejnené aj datasety bez licencie alebo s nevhodnou, obmedzujúcou licenciou. 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 v štatistike na portáli data.gov.sk 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 (napr. XML, xml, xml.net, xslx, .xslx, PDF, pdf, HTML, html, htm, CSV, csv, csw) alebo v jednom tagu sú uvedené viaceré formáty s čiarkou alebo s bodkou. 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 formáty sú zaradené aj hypertextové odkazy na iné sídlo mimo data.gov.sk (http, url), tie by mohli byť odlišené od súborov na data.gov.sk.
6. Existuje zrozumiteľné hodnotenie otvorených dát hviezdčkami²¹, známe na Slovensku aj vo svete. Kvalita poskytovaného datasetu je v tomto význame zadefinovaná vo výnose MF SR č. 55/2014 Z. z. (Výnos o štandardoch pre informačné systémy verejnej správy)²², § 51 až § 53. Minimálna požadovaná úroveň pre otvorené dáta je úroveň 3 hviezdčiek. Portál data.gov.sk by mohol automaticky hodnotiť zverejnené datasety hviezdčkami na základe platného výnosu, 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. Zvážiť striktné požiadavky na zverejnené datasety – portál data.gov.sk by mohol umožniť zverejnenie len tým datasetom, ktoré splnia aspoň 3 hviezdčky – čo by bolo v súlade s výnosom MF SR č. 55/2014 Z. z. (Výnos o štandardoch pre informačné systémy verejnej správy)²³, § 51 až § 53.

²⁰napr. <https://data.gov.sk/dataset/vestnik-verejneho-obstaravania-20195>

²¹<https://5stardata.info/cs/>

²²<https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2014/55/20190601>

²³<https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2014/55/20190601>

8. Zvážiť sprístupnenie hodnotenia a komentovania datasetov aj pre neprihlásených používateľov portálu data.gov.sk. Dôvod: prehľadnosť pre používateľov, spätná väzba pre poskytovateľov datasetov.
9. Verejnosť a viacerí dátoví kurátori organizácií by chceli vidieť štatistiku prístupov na jednotlivé datasety. Tým by sa tiež zhodnotil dopyt verejnosti po najžiadanejších datasetoch. Avšak vytvoriť objektívnu štatistiku prístupov je ťažké, lebo niektoré jednotlivé prístupy môžu byť duplicitné, niektoré viaceré prístupy agregované do jedného a niektoré nezaznamenané (ak dataset je zverejnený mimo portálu data.gov.sk). Portál by mohol pri datasetoch zobrazovať napr. jednoduchú grafickú štatistiku prístupov podľa času.
10. Portál data.gov.sk by mohol zobrazovať štatistiku využitia jednotlivých podaní²⁴ 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 dataset²⁵. Také nedostatky treba odstrániť, aby portál data.gov.sk bol v súlade s odporúčaniami WCAG 2.0, ktoré sú obsiahnuté aj vo výnose MF SR č. 55/2014 Z. z. (Výnos o štandardoch pre informačné systémy verejnej správy)²⁶, § 14 až § 17.
12. Portál data.gov.sk by mohol byť doplnený o ďalšie štatistiky alebo analytické prehľady (všetky podľa tejto analýzy). Ďalšia analýza by bola automatizovaná, objektívnejšia, bez rozdielov v metodike medzi rokmi.

Odporúčanie pre povinné osoby – poskytovateľov datasetov

*Metodická príručka pre povinné osoby – ako zverejňovať otvorené dáta (2015)*²⁷ 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.

Všetky uvedené odporúčania by bolo vhodné zapracovať do dokumentov, stratégií, akčných plánov ÚPPVII alebo iných zodpovedných inštitúcií. Úrad splnomocnenca vlády SR pre rozvoj občianskej spoločnosti v týchto mesiacoch zastrešuje participatívny proces tvorby nového *Akčného plánu Iniciatívy pre otvorené vládnutie na roky 2019 – 2021*, kde je tiež priestor pre niektoré úlohy.

²⁴<https://data.gov.sk/published-services>

²⁵napr. <https://data.gov.sk/dataset/pocet-rozhodnuti-podani-a-notifikacii-zaslanych-cez-upvs>

²⁶<https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2014/55/20190601>

²⁷https://www.slovensko.sk/_img/CMS4/Navody/Metodicka_prirucka_pre_povinne_osoby.pdf