

Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát 2021

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, 30.06.2021



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 API	9
1.6 Aktualizácie datasetov	9
1.7 Metadáta k datasetom	9
1.8 Používatelia datasetov	10
2 Obsah zverejňovaných datasetov	12
2.1 Publikačné minimum	12
2.2 Prieskum dopytu verejnosti po najžiadanejších datasetoch a API	13
2.3 Užitočnosť otvorených dát pre používateľa	14
2.4 Rámec na hodnotenie kvality otvorených dát a analýzu dátových portálov	14
3 Kvalitatívna analýza	15
3.1 Hodnotenie datasetov	15
3.2 Global open data index	17
3.3 Open data barometer	17
3.4 Open data watch	17
3.5 Index digitálnej ekonomiky a spoločnosti (DESI)	18
3.6 Open data maturity index	18
3.7 Nástroje	19
3.8 Aplikácie	19
3.9 API	20
Záver	21

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, 2020. 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 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 *Ministerstvo investícií, regionálneho rozvoja a informatizácie SR* (MIRRI), ktoré 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.mirri.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, s nejakou priradenou licenciou, pričom výber formátov a licencií môže byť duplicitný alebo nesprávny. Dataset s daným názvom a od danej organizácie sa považuje aktuálne za jeden, bez ohľadu na zdrojové dáta, obdobia, formát datasetu. Niektoré datasety sú zverejňované osobitne za rôzne obdobia (napr. denne, mesačne, ročne). 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 dátumom: 27.08.2013, 08.03.2015, 31.03.2017, 30.06.2018, 25.06.2019, 20.05.2020, 17.06.2021. Portál data.gov.sk v záložke Štatistiky uvádza počet organizácií, sumárny počet datasetov³.

Zoznam ústredných orgánov štátnej správy je prevzatý z portálu www.slovensko.sk⁴, 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.sk⁵, 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. Portál data.gov.sk v záložke Štatistiky uvádza počet organizácií 91 a celkový počet datasetov 2 433 (líši sa od súčtu v tabuľke).

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

⁴ <https://www.slovensko.sk/sk/agendy/agenda/ministerstva-a-ostatne-ustredne1>

⁵ <https://www.slovensko.sk/sk/institucie>

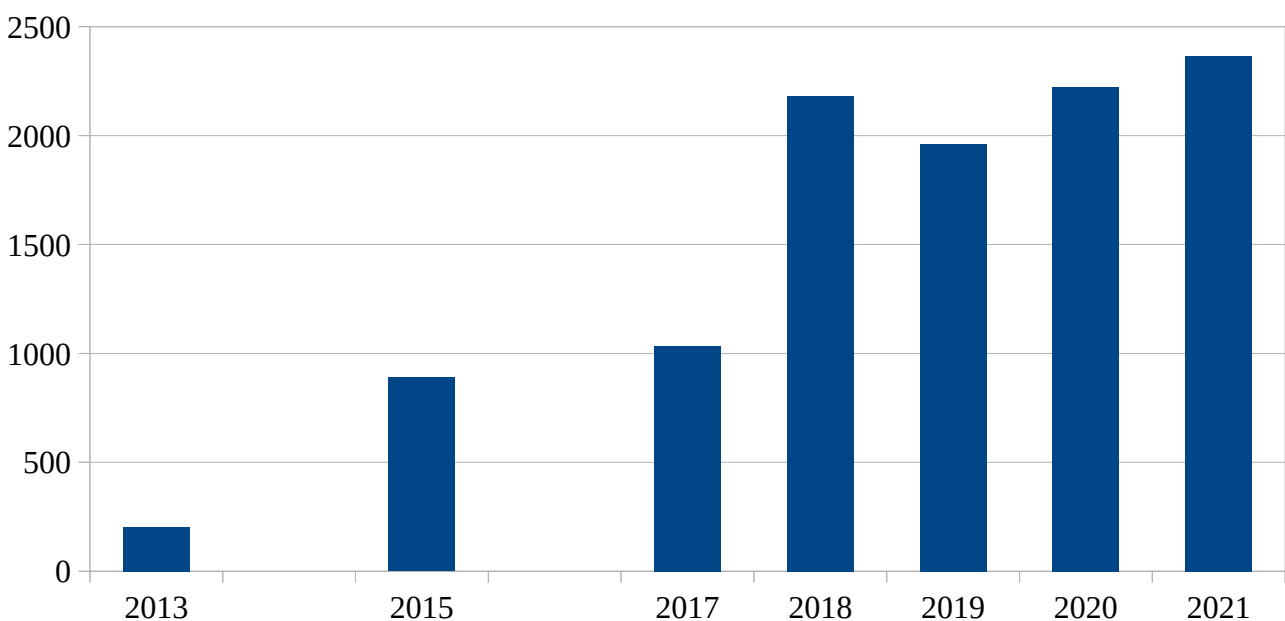
Tabuľka 1: Počet zverejnených datasetov podľa organizácií a rokov

Organizácia	2013	2015	2017	2018	2019	2020	2021
Ministerstvo dopravy a výstavby SR	8	10	10	32	66	60	77
Ministerstvo financií SR	11	22	22	20	13	14	15
Ministerstvo hospodárstva SR	7	8	9	17	17	34	37
Ministerstvo investícií, regionálneho rozvoja a informatizácie SR	0	23	12	36	39	36	58
Ministerstvo kultúry SR	10	11	11	35	25	26	26
Ministerstvo obrany SR	3	4	4	19	8	8	8
Ministerstvo pôdohospodárstva a rozvoja vidieka SR	65	68	67	209	457	469	475
Ministerstvo práce, sociálnych vecí a rodiny SR	5	5	5	5	13	13	13
Ministerstvo spravodlivosti SR	5	8	8	17	16	17	21
Ministerstvo školstva, vedy, výskumu a športu SR	5	24	24	34	26	26	23
Ministerstvo vnútra SR	24	40	41	253	61	62	62
Ministerstvo zahraničných vecí a európskych záležitostí SR	14	14	16	25	25	11	12
Ministerstvo zdravotníctva SR	9	18	95	184	262	381	408
Ministerstvo životného prostredia SR	9	9	22	41	25	42	52
Národný bezpečnostný úrad				4	23	23	23
Protimonopolný úrad SR				7	7	10	10
Správa štátnych hmotných rezerv SR				5	7	10	14
Štatistický úrad SR	9	586	606	1 101	665	745	764
Úrad geodézie, kartografie a katastra SR	2	2	2	2	2	5	19
Úrad jadrového dozoru SR				6	7	7	7
Úrad pre normalizáciu, metrológiu a skúšobníctvo SR				3	7	7	7
Úrad pre verejné obstarávanie	7	7	46	16	79	90	103
Úrad priemyselného vlastníctva SR				19	8	9	10
Úrad vlády SR	7	11	11	13	15	18	22
Generálna prokuratúra SR				8	8	8	8
Kancelária prezidenta SR							
Kancelária súdnej rady SR							
Kancelária verejného ochrancu práv							
Národná rada SR					0	10	10
Najvyšší kontrolný úrad SR					2	10	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	71
Spolu	204	889	1 034	2 181	1 962	2 222	2 365

Poznámka: Ministerstvo investícií, regionálneho rozvoja a informatizácie SR (MIRRI) je nástupníckym orgánom po Úrade podpredsedu vlády SR pre investície a informatizáciu (ÚPVII) od roku 2020.

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 metadátam 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.

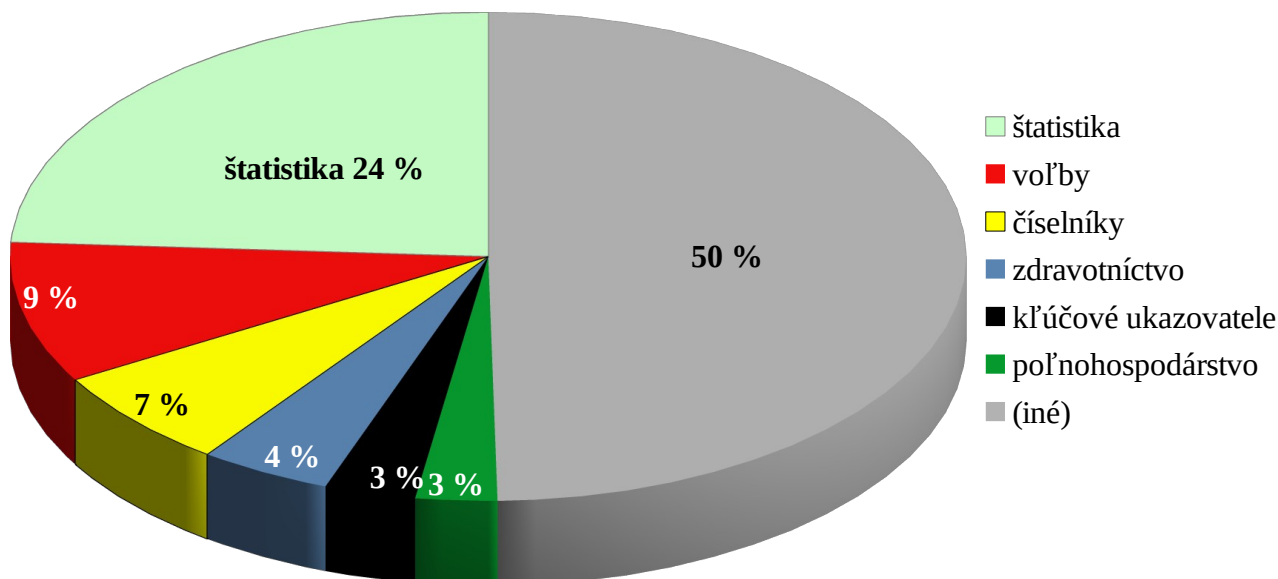
1.2 Tagy (kľúčové slová)

Portál data.gov.sk v záložke Datasety umožňuje filtrovať datsety podľa tagu (kľúčového slova) pre ľahšie vyhľadávanie⁶, 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).

Oproti minulému roku nie je v poradí a významnosti tagov zásadný rozdiel.

⁶ https://data.gov.sk/dataset? tags_limit=0

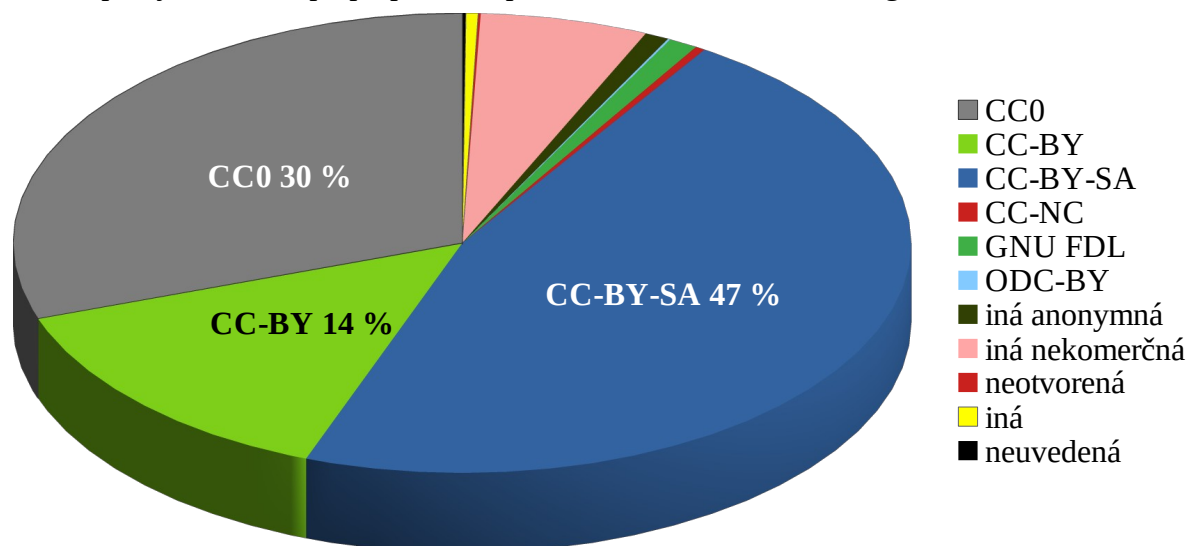


Graf 2: Počet datasetov podľa tagov (kľúčových slov)

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.



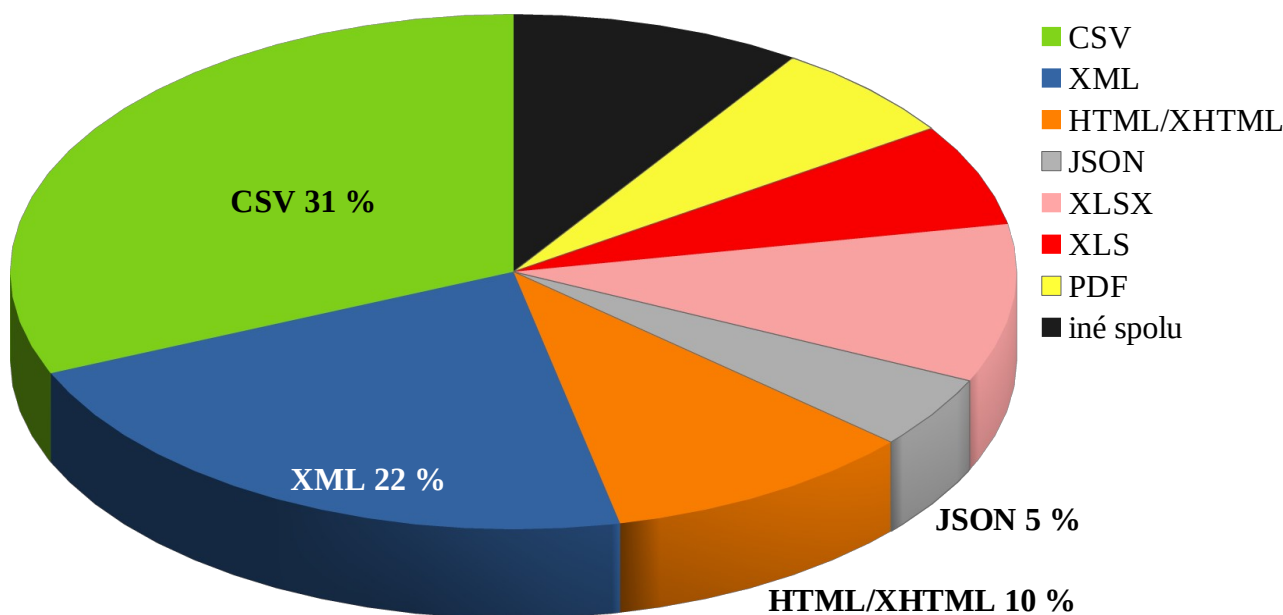
Graf 3: Typy licencií

Poradie najpoužívanejších licencií CC-BY-SA, CC0, CC-BY sa medzi rokmi nemení. 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 štandardmi⁷. Odporúčané licencie sú CC0, CC-BY. Poskytovateľ datasetov, ktorý neuvádza licenciu, je Úrad vlády SR.

⁷ 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 až 2021 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.⁸

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, xls“, „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.

⁸ <https://www.data.gov/developers/blog/primer-machine-readability-online-documents-and-data>

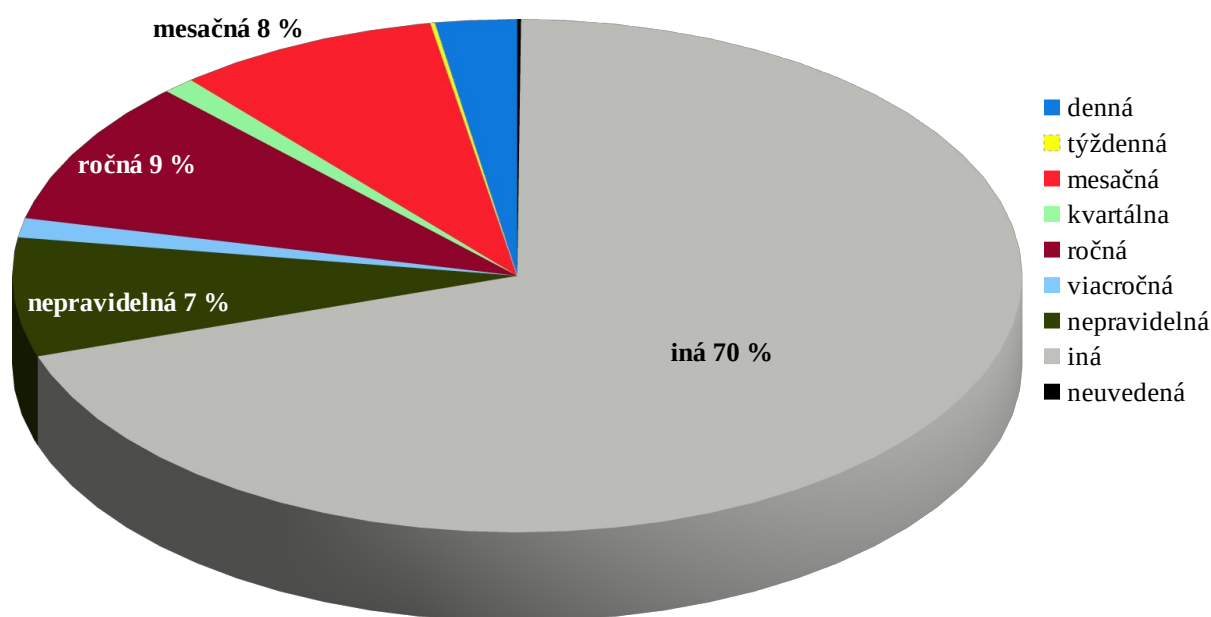
1.5 API

Strojové spracovanie dát môže byť vykonané na jednorázovo (ručne) stiahnutom datasete, na automaticky sťahovanom datasete, alebo automaticky prístupom na dátový zdroj cez API⁹.

Portál data.gov.sk nesprístupňuje jednotlivé API ku datasetom dostatočne, nevieme ľahko vyhodnotiť zverejňované datasety podľa tohto kritéria.

1.6 Aktualizácie datasetov

Datasety majú uvedený dátum zverejnenia datasetu, dátum zverejnenia dátového zdroja, dátum poslednej aktualizácie dátového zdroja a pravidelnosť aktualizácie.



Graf 5: Uvedená pravidelnosť aktualizácie datasetov

Pravidelnosť aktualizácie datasetov sme zistili zo zoznamu datasetov od prevádzkovateľa (NASES)¹⁰. Daný dataset môže byť v zozname uvedený viackrát, ak je zverejnený za rôzne zdrojové dáta a rôzne obdobia. Celkový počet datasetov sa potom líši od predošlých kapitol, uvádzame percentuálny prepočet. Je zbytočné analyzovať aktuálnosť datasetov, teda či daný dataset bol aktualizovaný podľa uvedenej pravidelnosti, pretože väčšina datasetov má uvedenú pravidelnosť aktualizácie: inú, nepravidelnú, neuvedenú – to sú nevhodné hodnoty.

1.7 Metadáta k datasetom

Dostupné metadáta k datasetom na portáli data.gov.sk alebo na inom webovom sídle majú rôzny rozsah a rôznu kvalitu (jednoznačnosť, správnosť, aktuálnosť).

⁹ https://sk.wikipedia.org/wiki/Application_programming_interface

¹⁰ Súbor: exportMod24052021.xlsx (24.05.2021)

V nasledujúcich analýzach by sa mohli (automaticky) spracúvať ďalšie metadáta datasetov, dostupné cez API alebo cez popisnú webovú stránku datasetu s lepšou sémantikou, ktoré na portáli data.gov.sk chýbajú, 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.

1.8 Používatelia datasetov

Na portáli data.gov.sk v záložke Štatistiky sa uvádza počet registrovaných používateľov portálu 5 276, čo je zvýšenie o 17 % oproti roku 2020, podobne aj pred rokom. Keďže datasety na portáli sú verejne prístupné, používateľ nemusí byť registrovaný, aby ich mohol prezeráť. Registrovaný používateľ dostáva 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). Pri niektorých službách je odkaz na z jedného portálu na druhý, pri niektorých odkaz chýba a používateľ musí viac hľadať. Portál data.gov.sk informuje o týchto poskytovaných službách¹¹:

- 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.

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

Verejnosť môže využiť nasledovné podania:

- Žiadosť o prístup k informáciám na účely opakovaného použitia podľa zákona o slobodnom prístupe k informáciám.¹²
- 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 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.*

Príprava publikačného minima je v omeškaní zhruba rok. V predošlých rokoch sa uskutočnili diskusie, je niekoľko prvotných návrhov alebo vzorov, aký obsah by publikačné minimum malo mať, či už pre ústredné orgány štátnej správy, verejnú správu všeobecne, alebo pre územnú samosprávu.

Smernica EP č. 2019/1024¹³ 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.

Z nasledujúcej kapitoly 2.2 o prieskume dopytu verejnosti po najžiadanejších datasetoch a API vyplýva, že Slovensko zaostáva vo zverejňovaní datasetov s vysokou hodnotou v oblastiach: Geopriestorové údaje (kataster), Spoločnosti a vlastníctvo spoločností (register právnických osôb), Mobilita (cestovné poriadky, polohy vlakov).

Dátová ekonomika (ekonomika postavená na dátach) sa opiera najmä o opakované použitie otvorených dát verejného sektora, čím sa vytvára ekonomická hodnota a úžitok pre spoločnosť.¹⁴ Slovensko tento potenciál nevyužíva dostatočne, čo aj vyplýva z medzinárodného hodnotenia v kapitolách 3.5 a 3.6.

Portál európskych údajov (europeandataportal.eu)¹⁵ 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.

¹³ https://datalab.digital/wp-content/uploads/Smernica2019_1024OPENDATA.pdf

¹⁴ <https://digital-strategy.ec.europa.eu/en/policies/open-data>

¹⁵ <https://www.europeandataportal.eu/data/datasets?locale=sk>

*Open Knowledge Foundation Global Open Data Index (OKFN GODI)*¹⁶ zatrieduje datasety do týchto kategórií: 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.

Ministerstvo investícií, regionálneho rozvoja a informatizácie SR (MIRRI) prostredníctvom dátovej kancelárie DATALAB¹⁷ participatívne pripravuje publikačné minimum pre štátnu správu aj pre samosprávu, so správne nastavenou štruktúrou, čo si vyžiada čistenie existujúcich dát.

Publikačné minimum orgánu štátnej správy (PUBMIN-CENTRAL-GOV)¹⁸ sa v roku 2021 priebežne špecifikuje, s predpokladaným termínom platnosti a účinnosti od 1.1.2022. To súvisí prípravou projektu Open Data 2.0¹⁹ podľa štandardu *W3C Data Catalog Vocabulary (DCAT)* 2²⁰. Nová verzia portálu data.gov.sk má fungovať aj ako robot, ktorý prehľadá webové sídla povinných osôb (aj samospráv) alebo iné portály a keď nájde vhodný dataset na zverejnenie, automaticky vytvorí prepojenie na tento dataset a doplní metadáta. Publikačné minimum samosprávy (MINPUB-LOCALGOV)²¹ vzniká v spolupráci dátovej kancelárie MIRRI, samosprávy a občianskeho združenia Alvaria²².

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 všetkých aktérov.

Úrad splnomocnenca vlády SR pre rozvoj občianskej spoločnosti každoročne od roku 2017 vykonáva prieskum dopytu verejnosti po najžiadanejších datasetoch a API.²³ 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.

Výsledkom priebežného prieskumu dopytu verejnosti po najžiadanejších datasetoch a API je zverejnená tabuľka²⁴, ktorá sa priebežne dopĺňa a aktualizuje. Niektoré datasety z prieskumu sú zverejnené na portáli data.gov.sk alebo na inom portáli, niektoré čiastočne, niektoré nie sú. Tabuľka je príliš veľká na zobrazenie vcelku na jednej strane, preto uvádzame ako príklad niekoľko najčastejšie dopytovaných datasetov, ktoré ešte neboli zverejnené alebo boli čiastočne zverejnené.

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

17 <https://datalab.community>

18 <https://wiki.vicempremier.gov.sk/pages/viewpage.action?pageId=64716866>

<https://metais.vicempremier.gov.sk/publicspace?pageId=64716866>

19 <https://datalab.digital/projekt-open-data-2-0/>

20 <https://www.w3.org/TR/vocab-dcat-2/>

21 <https://metais.vicempremier.gov.sk/publicspace?pageId=67145883>

22 <https://www.alvaria.sk>

23 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.

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

Tabuľka 2: Výber najžiadanejších datasetov – nezverejnených alebo čiastočne zverejnených

Organizácia	Stručný popis dát (sumarizácia)
Ministerstvo dopravy a výstavby SR	Cestovné poriadky verejnej dopravy samosprávnych krajov
	Aktuálne polohy a informácie o vlakoch (ŽSR/ŽSSK)
Ministerstvo financií SR	Rozpočtový informačný systém (RIS) rôznych organizácií
Ministerstvo spravodlivosti SR	Právne predpisy v strojovo spracovateľnej forme, vrátane histórie
	Legislatívny proces v strojovo spracovateľnej forme
Štatistický úrad SR	Register právnických osôb, register ekonomických subjektov
Úrad geodézie, kartografie a katastra SR	Katastrálne údaje vo vektorovom formáte, digitálny model reliéfu
	Katastrálne údaje o listoch vlastníctva, pozemkoch, stavbách, bytoch

Ministerstvo investícií, regionálneho rozvoja a informatizácie SR zverejnilo dataset *MIRRI Najžiadanejšie otvorené údaje (2021-05-11)*²⁵, v ktorom je málo záznamov.

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

2.3 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 %

2.4 Rámec na hodnotenie kvality otvorených dát a analýzu dátových portálov

Viacere štáty riešia podobné problémy v oblasti otvorených dát ako Slovensko. Môžeme spojiť sily a využiť rovnaký spôsob hodnotenia kvality otvorených dát a analýzy dátových portálov.

*Open Data and the Status Quo – A Fine-Grained Evaluation Framework for Open Data Quality and an Analysis of Open Data portals in Germany*²⁷ je článok, ktorý opisuje rámec hodnotenia kvality otvorených dát a metadát na portáloch otvorených dát. Navrhované kritériá pre hodnotenie kvality otvorených dát: kľúčové dáta, témy, aktuálnosť, pokrytie územia, jedinečnosť, interoperabilita, právna otvorenosť, dostupnosť a prístupnosť, úplnosť podľa DCAT.

25 <https://data.gov.sk/dataset/mirri-naziadanejsie-otvorene-udaje-ovm-2021-05-11>

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

27 <https://arxiv.org/abs/2106.09590>

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.

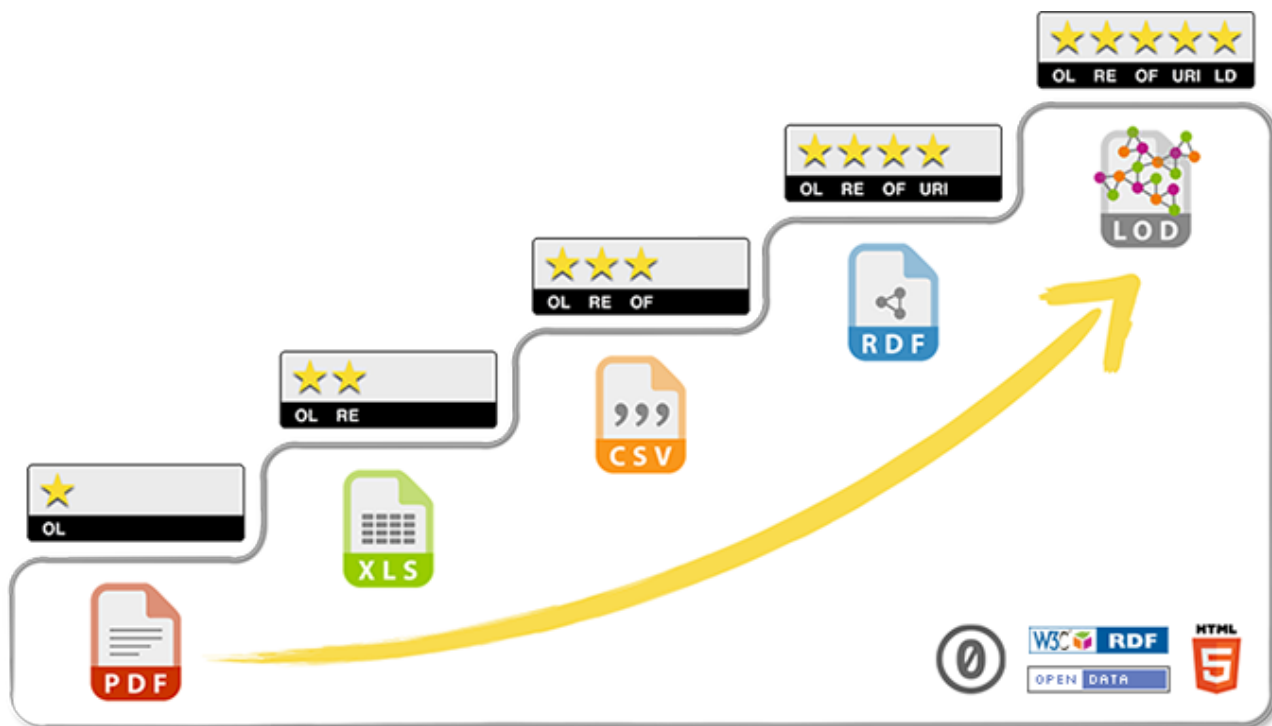
Špeciálnym typom datasetu je katalóg (zoznam) zverejňovaných datasetov danej organizácie.

3.1 Hodnotenie datasetov

Tim Berners-Lee, vynálezca webu a iniciátor prepojených dát, navrhol hodnotenie otvorených dát hviezdami²⁸:

- * 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 č. 78/2020 Z. z. (Vyhláška o štandardoch pre informačné technológie verejnej správy)²⁹, § 38 až § 40.



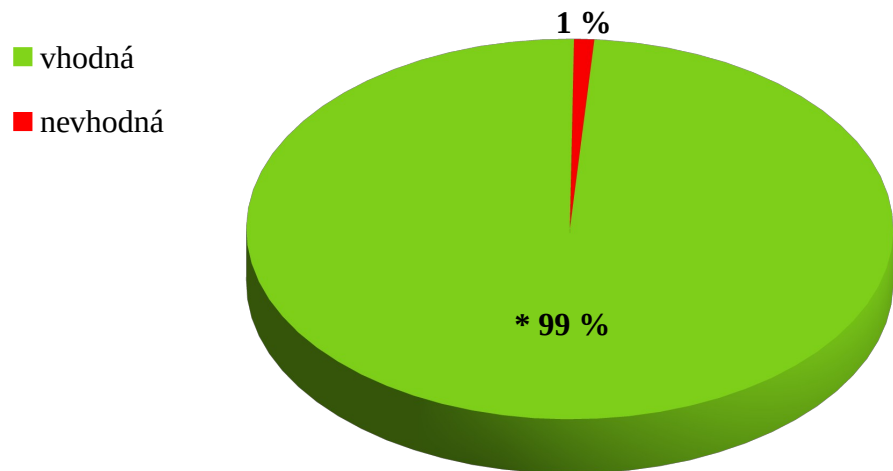
Obrázok 2: Hodnotenie otvorených dát hviezdami

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

²⁸ <https://5stardata.info/cs/>

²⁹ <https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2020/78/20200501>

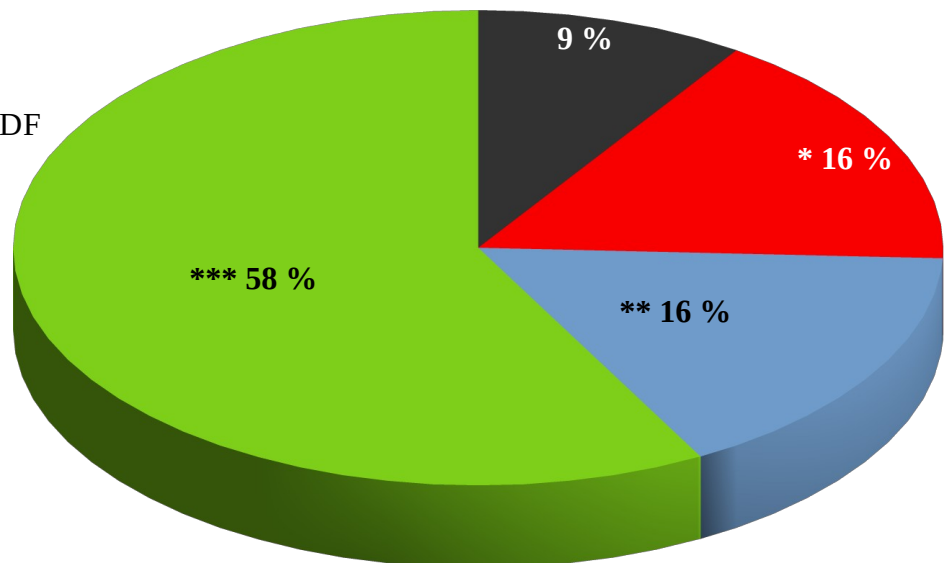
Ohodnotenie výberu licencie hviezdíčkami je jednoduché. Všetky typy licencií uvedené na portáli data.gov.sk, ktoré boli použité v grafe 3 (licencie datasetov), sme rozdelili na dve skupiny – vhodná a nevhodná. Vhodné alebo štandardmi odporúčané sú (otvorené) verejné licencie CC0, CC-BY, CC-BY-SA, GNU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



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

Vzhľadom na veľmi malý podiel nevhodných licencií možno usúdiť, že prienik hodnotenia licencií a formátov bude takmer rovnaký ako graf 7.

Formát RDF, ktorý poskytuje ontológiu pre metadáta ku datasetu, je na portáli data.gov.sk len pri 11 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 použitie referencovateľných identifikátorov v dátach. Možno usúdiť, že zrejme všetky 3-hviezdičkové datasety zverejnené na portáli data.gov.sk sú aj 4-hviezdičkové.

Úroveň 5 hviezdíčiek splňajú tie datasety, ktoré majú 4 hviezdíčky a obsahujú v dátach prepojenia referencovateľnými identifikátormi na iné datasety. Toto nevieme dostatočne rýchlo sumárne vyhodnotiť z dostupných informácií o datasetoch, iba pri jednotlivých datasetoch.

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

V roku 2021 Ministerstvo investícií, regionálneho rozvoja a informatizácie SR zverejnilo päť datasetov s formátom RDF, niektoré aj s prepojenými dátami.

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 2020/21³⁰ na 10. mieste so 78 % v rebríčku 187 hodnotených krajín (Singapur na 1. mieste s 92 %). Oproti predošlému obdobiu 2018 sa Slovensko zlepšilo. Posudzujú sa oficiálne zverejňované štatistiky štátu v oblastiach: sociálna, ekonomická, životné prostredie. Prehľad o Slovensku za rok 2020³¹ je na základe dát Štatistického úradu SR³².

30 <https://odin.opendatawatch.com/Report/annualReport2020#sec5>

31 <https://odin.opendatawatch.com/Report/Reports>

32 <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) 2020³³ patrí Slovensku 22. miesto v rebríčku 28 krajín EÚ, s poklesom o 1 miesto oproti roku 2019, rovnako v roku 2019 pokles o 1 miesto oproti roku 2018. Slovensko je v hodnotení DESI pod priemerom EÚ s podpriemerným rastom za roky 2015 – 2020, čo znamená, že dlhodobo zaostáva, umiestňuje sa v najslabšej tretine členských štátov EÚ.

Index DESI hodnotí členské štáty EÚ vo výkonnosti podľa stratégie pre jednotný digitálny trh a digitálnu konkurencieschopnosť, s prihliadnutím na program Digitálna Európa. Ukazovatele sú rozdelené do piatich dimenzií: konektivita, ľudské zdroje (digitálne zručnosti), používanie internetu, integrácia digitálnej technológie, verejné digitálne služby. Najlepšie hodnotenie má Fínsko a najrýchlejší rast má Írsko.

Najväčší pokles Slovensko zaznamenalo v ukazovateli otvorených dát (dimenzia verejné digitálne služby). V ostatných ukazovateľoch nie je vidno výrazné zaostávanie alebo napredovanie oproti celkovému hodnoteniu.

V medzinárodnom porovnaní³⁴ v niektorých dimenziách DESI sú najlepšie tieto štáty (lepšie ako EÚ): Japonsko, Island v konektivitě. USA, Austrália v ľudskom kapitáli (digitálnych zručnostiach). Island, Nórsko v používaní internetu. Švajčiarsko v integrácii digitálnej technológie. Južná Kórea, USA vo verejných digitálnych službách.

Čo robí Slovensko pre zlepšenie v indexe DESI? Pripravuje sa Stratégia a akčný plán na zlepšenie postavenia Slovenska v indexe DESI do roku 2025.³⁵

3.6 Open data maturity index

<https://data.europa.eu/en/impact-studies/open-data-maturity>

Slovensko sa umiestnilo v poslednom sledovanom období 2020³⁶ v najslabšej skupine sledovaných európskych štátov (začiatovníci) zo štyroch skupín (lídri trendu, rýchli nasledovníci trendu, nasledovníci, začiatovníci). Dánsko je na 1. mieste. Metodika je založená na posúdení 4 dimenzií otvorených dát: politika, portál, dopad, kvalita.

Podrobnejší pohľad na Slovensko 2020³⁷ ukazuje, že prepád nastal v roku 2018 a prehýbil sa v roku 2019. Najlepší výsledok je v dimenzii politika, na priemernej úrovni v porovnaní s EÚ. 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: použitie portálu, poskytovanie dát a udržateľnosť portálu. Dimenzia kvalita má zlé hodnotenie podľa kritérií: monitoring a opatrenia, aktuálnosť

33 https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=67086

34 https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=72352

35 <https://www.slov-lex.sk/legislativne-procesy/SK/LP/2021/126>, <https://platforma.slovensko.digital/t/strategia-a-akcny-plan-pre-zlepsenie-postavenia-sr-v-indexe-desi-do-roku-2024/7261>

36 <https://data.europa.eu/en/dashboard/2020>

37 https://data.europa.eu/sites/default/files/country-factsheet_slovakia_2020.pdf

a úplnosť, súlad so štandardom DCAP. Hodnotenie sa vytvára na základe dotazníka, najnovší bude k polroku 2021.³⁸

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).

Čo robí Slovensko pre zlepšenie? Pripravuje sa projekt nového portálu Open Data 2.0, vyhovujúci štandardu DCAT, s meraním kvality datasetov.

3.7 Nástroje

Na portáli data.gov.sk v záložke Nástroje sú dva nástroje. 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.

MOD_EDEM_PA

Podporná aplikácia pre povinné osoby.

3.8 Aplikácie

Na portáli data.gov.sk v záložke Aplikácie je dostupných 5 aplikácií, čo je katastrofálny stav po mnohých rokoch fungovania portálu, aj keď ide o malé zlepšenie oproti minulému roku. Na tomto mieste by mali byť zverejnené desiatky až stovky rôznych používateľských aplikácií na prácu so zverejnenými dátami (datasetmi). Prekážkou pri zverejňovaní aplikácií je neprívetivá registrácia, povinné použitie eID, kvôli problémom so samotným použitím eID a portálom slovensko.sk.

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).

38 <https://platforma.slovensko.digital/t/index-dospelosti-otvorených-udajov-odmi-sk-pripomienkovanie/7692>

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

Autoform

Automatické vyplňanie kontaktných a fakturačných údajov o právnických osobách.

Služby Slovensko.Digital, s.r.o., 2019, <https://ekosystem.slovensko.digital/sluzby/autoform>

Notifikátor verejných obstarávaní

Nastavenie emailových notifikácií pre novo vyhlásené verejné obstarávania vo vestníku.

UVOstat.sk, 2020, <https://www.uvostat.sk/vyhlasene-obstaravania>

- **Aplikácie na iných portáloch**

Zverejnenie aplikácie na inom portáli nevyžaduje použitie eID, je jednoduchšie. Dobrým príkladom je komunita OpenData.sk³⁹ a Utopia.sk⁴⁰, ktorá má podstatne viac zverejnených aplikácií:

<https://utopia.sk/wiki/display/opendata/SK+a+CZ+opendata+aplikacie+a+projekty>

3.9 API

Portál data.gov.sk poskytuje CKAN API⁴¹ v. 2.2.3 pre prácu s datasetmi a metadátami datasetov. API⁴² môže slúžiť na automatizované spracovanie dát a metadát pre používateľov, vytváranie webových stránok s otvorenými dátami, vytváranie portálov, aj automatické generovanie analýzy datasetov.

Originálna dokumentácia CKAN (*Comprehensive Knowledge Archive Network*)⁴³ je dostupná už v novšej verzii 2.9.3 a nástroj CKAN API je v aktuálnej verzii 3.

39 <https://opendata.sk>

40 <https://utopia.sk>

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

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

43 <https://docs.ckan.org/en/2.9/>
<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. Zvážiť informovanie o prieskume aj cez notifikácie portálu data.gov.sk pre registrovaných používateľov.
- **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 2020* pre Slovensko (kap. 3.6). Preto je užitočné realizovať projekt Open Data 2.0.
- **Rozšírenie záberu dát** – vzhľadom na hodnotenie otvorených dát hviezdikami (kap. 3.1), na *Stratégiu digitálnej transformácie Slovenska 2030*⁴⁴, na program *Digitálna Európa 2021 – 2027*⁴⁵ 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 2020* 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**.

44 <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>

45 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 analýzách z roku 2019 a 2020. Pridávame odporúčania, ako ich vyriešiť, alebo ako lepšie navrhnúť nový portál s projektom Open Data 2.0:

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. Rozdiel môže byť spôsobený aj nejednoznačným zadefinovaním, čo je dataset, distribúcia datasetu, formát datasetu, dátový zdroj, URI na dátový zdroj.
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 správy⁴⁶ a k nim priradené podriadené organizácie⁴⁷. Napríklad Regionálne úrady verejného zdravotníctva priradené k Ministerstvu zdravotníctva SR. Tiež všetky územné samosprávy združiť do jedného zoznamu. Niektoré organizácie na portáli data.gov.sk majú nepresné alebo duplicitné názvy.
3. Portál data.gov.sk zobrazuje počty datasetov podľa organizácií v niektorých prípadoch agregovane, teda dataset s daným názvom a od danej organizácie je považovaný za jeden dataset, bez ohľadu na zdrojové dáta, obdobia, formát datasetu a v iných prípadoch osobitne za rôzne zdrojové dáta alebo rôzne obdobia. Toto je potrebné urobiť jednotne.
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. Ponuka licencií by mala byť definovaná v číselníku, aby nevznikali nesprávne označenia licencií.
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.net“, „xlsx“, „.xlsx“, „XLS“, „xml, xls“). Tieto formáty treba zjednotiť. Ponuka formátov by mala byť definovaná v číselníku, aby nevznikali nesprávne označenia formátov. Medzi formátmi môžu byť chybné uvedené prepojenia (api), alebo hypertextové odkazy na iné sídlo mimo data.gov.sk (url, http), tie by mohli byť odlíšené od súborov na data.gov.sk.
6. Existuje zrozumiteľné hodnotenie otvorených dát hviezdikami⁴⁸. Kvalita poskytovaného datasetu je v tomto význame zadefinovaná vo vyhláske ÚPVII č. 78/2020 Z. z. (Vyhláska

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

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

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

o štandardoch pre informačné technológie verejnej správy)⁴⁹, § 38 až § 40. Na zverejňovanie otvorených dát sa vzťahuje tiež Národná koncepcia informatizácie verejnej správy⁵⁰, Stratégia a akčný plán prístupnosti a používania otvorených údajov verejnej správy⁵¹. Minimálna požadovaná úroveň pre otvorené dáta je úroveň 3 hviezdíčiek. Portál data.gov.sk by mohol automaticky hodnotiť zverejnené datasety hviezdíčkami na základe platného predpisu (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 hviezdíčky – čo by bolo v súlade s platnými predpismi (vymenované v predošlom bode).
8. Portál data.gov.sk by mal ponúkať povinnej osobe výber pravidelnosti aktualizácie datasetu z číselníka s odporúčanými zmysluplnými hodnotami, pričom treba obmedziť hodnoty „iná“, „neuveďená“.
9. Vyzvať povinné osoby, za pomoci dátových kurátorov, aby odstránili nedostatky zverejnených datasetov s menej ako 3 hviezdíčkami a zmenili ich na datasety s aspoň 3 hviezdíčkami. Toto opatrenie pomôže zlepšiť hodnotenie Slovenska v medzinárodných indexoch.
10. Zaviesť prístupnosť 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.
11. 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. Portál by mohol pri datasetoch zobrazovať jednoduchú grafickú štatistiku prístupov na popisné HTML stránky datasetov podľa času, štatistiku stiahnutí jednotlivých datasetov, štatistiku prístupov sumárne za celý portál – a to aj vo forme datasetu.
12. Portál data.gov.sk by mohol zobrazovať štatistiku využitia jednotlivých podaní⁵² verejnosťou.
13. 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 dataset⁵³. 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 služieb⁵⁴.

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

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

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

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

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

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

14. 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.
15. Portál data.gov.sk by mal umožniť zverejnenie a vyhľadávanie API ku datasetom podľa rovnakých kritérií ako zverejňovanie a vyhľadávanie samotných datasetov. Táto funkcia portálu by pomohla podstatne zlepšiť hodnotenie Slovenska v medzinárodných indexoch podľa kritérií použiteľnosť portálu, dopad na ekonomiku a spoločnosť.
16. 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 a vyhľadávať aplikácie aj na iných portáloch⁵⁵.
17. V rámci projektu Open Data 2.0 prejsť na CKAN API v najnovšej verzii 3.
18. V rámci projektu Open Data 2.0 pridať funkciu robota, ktorý prehľadá webové sídla povinných osôb (aj samospráv) alebo iné portály a keď nájde vhodný dataset na zverejnenie, automaticky vytvorí prepojenie na tento dataset a doplní metadáta.
19. Zverejniť informačný sumárny dataset o všetkých zverejnených datasetoch pre uľahčenie analýz zverejňovania datasetov otvorených dát, ako je napr. táto analýza.
20. Katalogizovať datasety na portáli data.gov.sk a dátové zdroje pre rôzne použitia (hospodárstvo, zamestnanosť, vzdelávanie, životné prostredie atď.).
21. Doplniť novú sekciu portálu data.gov.sk, v ktorej by sa pravidelne zverejňovali novinky, články, blogy z problematiky otvorených dát, s príkladmi použitia otvorených dát, nástrojov, aplikácií. Týmto sa naplní úloha **Akčného plánu Iniciatívy pre otvorené vládnutie v SR na roky 2017 – 2019**, v kapitole Otvorené informácie:
 7. *Uskutočniť verejnú kampaň na propagáciu využitia datasetov zverejnených na portáli otvorených dát a na podporu vývoja inovácií. Termín: priebežne.*

55 <https://utopia.sk/wiki/display/opendata/SK+a+CZ+opendata+aplikacie+a+projekty>

Odporúčania pre dátovú kanceláriu MIRRI

1. Po vzore ČR⁵⁶ 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)*⁵⁷ obsahuje všetko potrebné pre správne zverejňovanie datasetov. Každá povinná osoba, u ktorej sa vyskytol nejaký nedostatok opísaný v tejto analýze, by mala postupovať podľa metodického príručky a napraviť nedostatky.

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

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