

Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2022

na základe úlohy B.10

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, 30.06.2022



Pre Úrad splnomocnenca vlády SR pre rozvoj občianskej spoločnosti spracoval
Lukáš Jankovič, OZ Alvaria, prvý draft sripomienkoval a doplnil **Martin Šechný**,
SOIT

Zoznam skratiek

API – aplikačné programové rozhranie

ÚSV ROS – Úrad splnomocnenca vlády SR pre rozvoj občianskej spoločnosti

MIRRI – Ministerstvo investícií, regionálneho rozvoja a informatizácie SR

OVM – orgán verejnej moci

PM – publikačné minimum

NASES – Národná agentúra pre sieťové a elektronické služby

1	Úvod	5
2	Kvantitatívna analýza	7
2.1	Počet zverejnených datasetov	7
2.2	Počet zverejnených datasetov podľa OVM	8
2.3	Tagy (kľúčové slová)	10
2.4	Licencie datasetov	11
2.5	Formáty datasetov	13
2.6	Aktualizácia datasetov	15
2.7	Zhrnutie	17
3	Kvalitatívna analýza	18
3.1	Hodnotenie datasetov podľa hviezdičiek	18
3.2	Najčastejšie problémy znižujúce použiteľnosť datasetov	20
3.3	Rebríčky	21
	Open data maturity index	21
	OECD OURdata Index - vládne otvorené dáta	22
	Open data watch	22
	Index digitálnej ekonomiky a spoločnosti (DESI)	23
	Ostatné rebríčky	23
3.4	Zhrnutie	23
4	Používanie portálu	25
4.1	Počet registrovaných používateľov portálu	25
4.2	Návštevníci podľa krajín	25
4.3	Počet zobrazení podľa rokov	26
4.4	Najpopulárnejšie datasety	27
4.5	Počet aplikácií	29
4.6	Zhrnutie	29
5	Štandardy a iniciatívy v oblasti otvorených dát	31
5.1	Publikačné minimum štátnej správy	31
5.2	Datasety s vysokou pridanou hodnotou	32
5.3	Prieskum dopytu verejnosti po najžiadanejších datasetoch a API	32
5.4	Národná koncepcia informatizácie verejnej správy	33
5.5	Národný projekt Otvorené údaje 2.0	34
5.6	Výzva Manažment údajov inštitúcie verejnej správy	34
5.7	Akčný plán Iniciatívy pre otvorené vládnutie na roky 2022-2024	35
5.8	Občianske združenie Alvaria	36
5.9	Zhrnutie	37

6	Záver	38
7	Odporúčania	39
7.1	Opakujúce sa odporúčania z poslednej analýzy	39
7.2	Nové odporúčania	41

1 Úvod

Slovenská republika sa v roku 2011 zapojila do medzinárodnej *Iniciatívy pre otvorené vládnutie*¹. Úrad splnomocnenca vlády SR pre rozvoj občianskej spoločnosti zastrešuje túto iniciatívu krátkodobými, spravidla 2 ročnými akčnými plánmi. Posledný schválený *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 správy vo svojej poslednej aktualizácii v roku 2021² a Stratégia a akčný plán sprístupnenia a používania otvorených údajov verejnej správy³.

Predchádzajúce analýzy zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát boli realizované za roky 2017, 2018, 2019, 2020, 2021. Táto analýza mapuje zverejnené **datasety ústredných orgánov štátnej správy** na portáli otvorených dát (data.gov.sk, ďalej len „portál“), 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 datasety územnej samosprávy.



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

¹ <https://www.opengovpartnership.org/>

² <https://www.mirri.gov.sk/sekcie/narodna-koncepcia-informatizacie-verejnej-spravy-nikvs/index.html>

³ https://www.slovensko.sk/_img/CMS4/strategia_otvorene_udaje.pdf



**Analýza zverejňovania datasetov
ústredných orgánov štátnej správy
na portáli otvorených dát za rok
2022**

a správy informačných technológií verejnej správy. Portál poskytuje zverejňovania datasetov a API tiež vyhľadávanie a filtrovanie datasetov podľa rôznych kritérií, návody a iné súvisiace informácie.

2 Kvantitatívna analýza

V tejto kapitole **sumarizujeme počty datasetov podľa viacerých kritérií**. Počty porovnávame po jednotlivých rokoch v období, v ktorom bola vypracovaná analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát (ďalej len „analýza“) v danom roku.

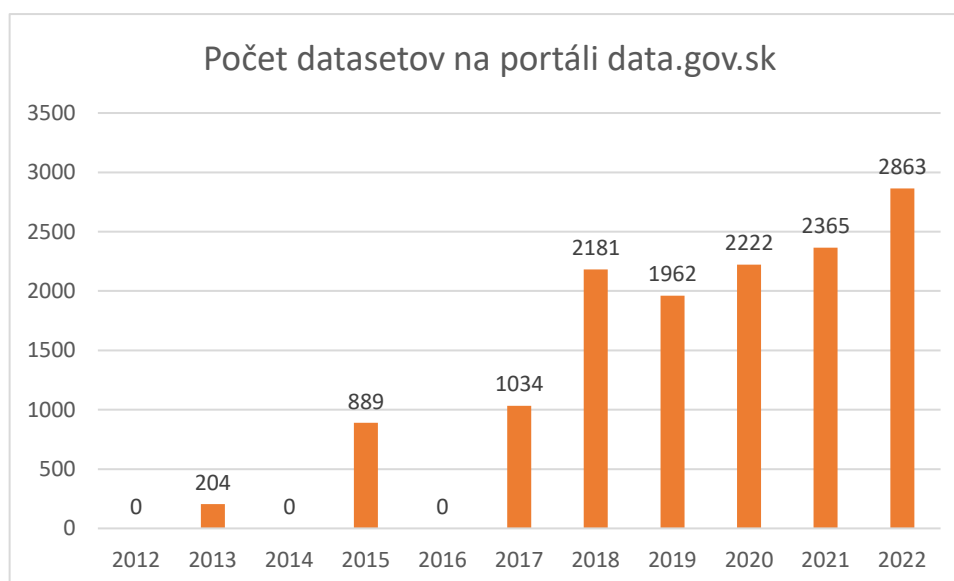
Metodika analýz z predchádzajúcich rokov sa môže líšiť, preto absolútne počty datasetov nemožno považovať za presné, sú skôr orientačné. Dáta boli zisťované k dátumom: 27.08.2013, 08.03.2015, 31.03.2017, 30.06.2018, 25.06.2019, 20.05.2020, 17.06.2021, 24.5.2022. Portál data.gov.sk v záložke Štatistiky uvádza počet organizácií, sumárny počet datasetov⁴.

Zdrojom údajov a štatistík v tejto kapitole je portál data.gov.sk a minulé analýzy, pokiaľ nie je uvedené inak.

2.1 Počet zverejnených datasetov

Aktuálny počet datasetov k 24.5.2022 bol zistený z portálu data.gov.sk. Počty datasetov za minulé roky sú prebraté z analýzy za rok 2021. **Počet datasetov postupne rastie, v rokoch 2020 a 2021 približne konštantným tempom.**

Graf 1

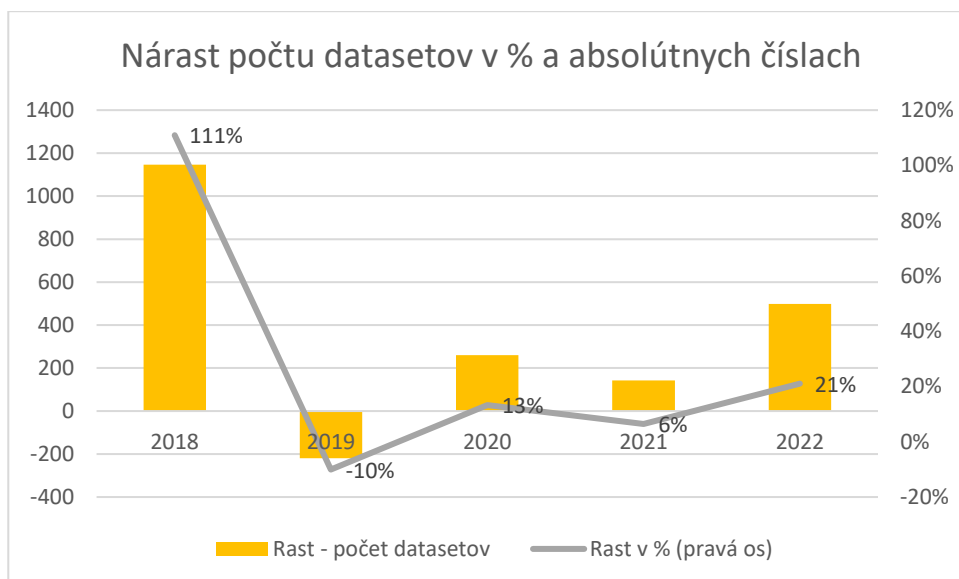


Pokiaľ sa pozrieme na percentuálny nárast v porovnaní s minulým rokom, zaznamenali sme **v roku 2022 rast 21 %**, čo je **najlepší výsledok za posledné 4 roky**. **Počet datasetov za posledný rok sa zvýšil o takmer 500**. Vyšší nárast bol zaznamenaný len v roku 2018 – 111 %.

Graf 2

⁴Sumárne počty v štatistike data.gov.sk (na obrázku) sa nezodujú so skutočným počtom datasetov po kliknutí.

Graf 2



Samotný nárast počtu datasetov nemusí nutne znamenať aj sprístupnenie nových dát, keďže je možné, že napríklad niektoré datasety sa rozdelili a vzniklo viacero datasetov. Je preto potrebné pozeráť sa aj na to ako sa zmenili datasety a či naozaj pribudli nové dáta.

2.2 Počet zverejnených datasetov podľa OVM

Nižšie ponúkame tabuľku počtu datasetov podľa orgánov verejnej moci (OVM). Zoznam ústredných orgánov štátnej správy je prevzatý z portálu www.slovensko.sk⁵, sú uvedené v tabuľke na bielych riadkoch. Do počtu datasetov organizácie sú započítané aj podriadené organizácie (napr. Ministerstvo pôdohospodárstva má až 16 podriadených organizácií). Okrem OVM zverejňuje datasety na portáli aj SK-NIC, a.s., súkromná spoločnosť, ktorá je správcom slovenskej internetovej domény sk. Datasety na portáli zverejňujú aj územné samosprávy.⁶

Tabuľka 1: Počet zverejnených datasetov na portáli podľa organizácií a rokov

OVM alebo názov organizácie	2013	2015	2017	2018	2019	2020	2021	2022	Rast
Generálna prokuratúra SR				8	8	8	8	8	0
Kancelária Národnej rady SR					0	10	10	10	0
Ministerstvo dopravy a výstavby SR	8	10	10	32	66	60	77	83	6
Ministerstvo financií SR	11	22	22	20	13	14	15	24	9
Ministerstvo hospodárstva SR	7	8	9	17	17	34	37	37	0
Ministerstvo investícií, regionálneho rozvoja a informatizácie Slovenskej republiky	0	23	12	36	39	36	58	100	42
Ministerstvo kultúry Slovenskej republiky	10	11	11	35	25	26	26	26	0

⁵https://www.slovensko.sk/sk/agendy/agenda/_ministerstva-a-ostatne-ustredne1

⁶ Ide o mestá Prešov, Trnava, Vranov nad Topľou, Nitriansky samosprávny kraj, obce Horné Strháre, Kolačkov

Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2022

Ministerstvo obrany SR	3	4	4	19	8	8	8	9	1
Ministerstvo pôdohospodárstva a rozvoja vidieka SR	65	68	67	209	457	469	475	491	16
Ministerstvo práce, sociálnych vecí a rodiny SR	5	5	5	5	13	13	13	13	0
Ministerstvo spravodlivosti SR	5	8	8	17	16	17	21	26	5
Ministerstvo školstva, vedy, výskumu a športu Slovenskej republiky	5	24	24	34	26	26	23	24	1
Ministerstvo vnútra SR	24	40	41	253	61	62	62	62	0
Ministerstvo zahraničných vecí a európskych záležitostí SR	14	14	16	25	25	11	12	12	0
Ministerstvo zdravotníctva SR	9	18	95	184	262	381	408	435	27
Ministerstvo životného prostredia SR	9	9	22	41	25	42	52	82	30
Najvyšší kontrolný úrad SR					2	10	10	32	22
Národný bezpečnostný úrad				4	23	23	23	23	0
Protimonopolný úrad SR				7	7	10	10	50	40
SK-NIC								1	1
Správa štátnych hmotných rezerv Slovenskej republiky				5	7	10	14	15	1
Štatistický úrad SR	9	586	606	1101	665	745	764	1023	259
Úrad geodézie, kartografie a katastra SR	2	2	2	2	2	5	19	19	0
Úrad jadrového dozoru Slovenskej republiky				6	7	7	7	12	5
Úrad pre normalizáciu, metrológiu a skúšobníctvo Slovenskej republiky				3	7	7	7	10	3
Úrad pre verejné obstarávanie	7	7	46	16	79	90	103	127	24
Úrad priemyselného vlastníctva Slovenskej republiky				19	8	9	10	15	5
Úrad vlády SR	7	11	11	13	15	18	22	22	0
Všeobecná zdravotná poisťovňa								1	1
Nezaradené	4	1	1	3	11				0
územná samospráva	0	18	22	67	68	71	71	67	-4
Spolu	204	889	1034	2181	1962	2232	2375	2869	494

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.

Portál data.gov.sk v záložke Štatistiky uvádza počet organizácií 94 (o 3 viac ako v roku 2021).

Najviac datasetov zverejňuje **Štatistický úrad SR**, ktorý prekonal hranicu 1000 datasetov, nasleduje **Ministerstvo pôdohospodárstva a rozvoja vidieka SR** a **Ministerstvo zdravotníctva SR** atakujú hranicu 500 datasetov. Viac ako 100 datasetov majú **Ministerstvo investícií, regionálneho rozvoja a informatizácie SR** a **Úrad pre verejné obstarávanie**.

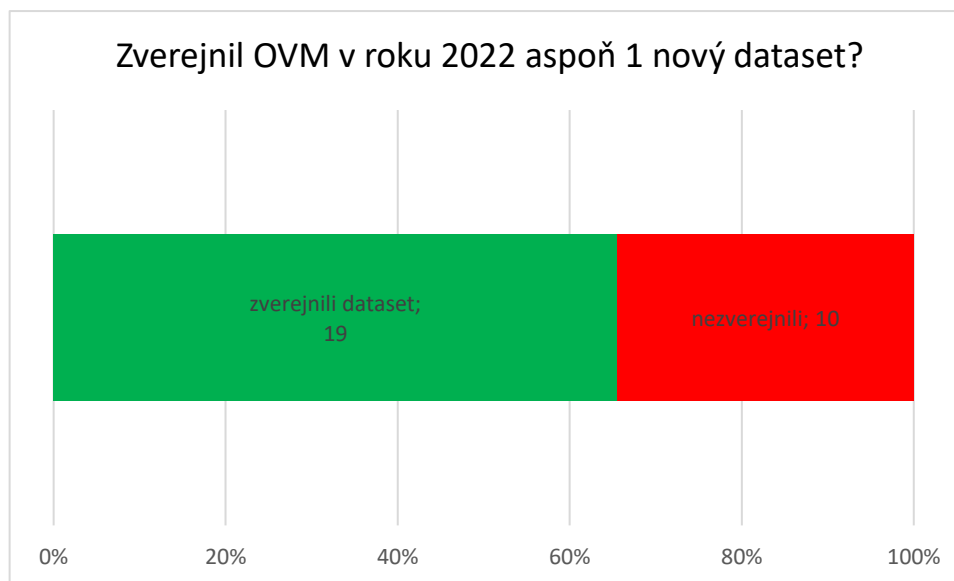
Najväčším skokanom za predchádzajúci rok je **Štatistický úrad SR**, ktorý pridal 259 datasetov, 42 datasetov pridalo **Ministerstvo investícií, regionálneho rozvoja a informatizácie SR**, 40 datasetov **Protimonopolný úrad SR**. Väčší počet datasetov ako 20 pridalo **Ministerstvo**

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

zdravotníctva SR, Najvyšší kontrolný úrad SR, Úrad pre verejné obstarávanie a Ministerstvo životného prostredia SR.

Počty datasetov pri organizáciách rástli alebo stagnovali, pokles zaznamenala len územná samospráva. Väčšina OVM však zverejnila aspoň 1 nový dataset.

Graf 3



Jeden z dôvodov, prečo je za OVM zverejnených málo datasetov (čo sa týka aj samosprávy) a sú zriedkavejšie aktualizované, je aj situácia, že informačný systém používaný OVM nezverejňuje automatizovane datasety, tie zverejňuje zamestnanec manuálne, pričom sa na portál otvorených dát prihlasuje svojim občianskym preukazom (musí mať oprávnenie na zastupovanie inštitúcie a pridelenú rolu na zverejňovanie datasetov).

2.3 Tagy (kľúčové slová)

Portál data.gov.sk v záložke Datasets umožňuje filtrovať datasety 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.

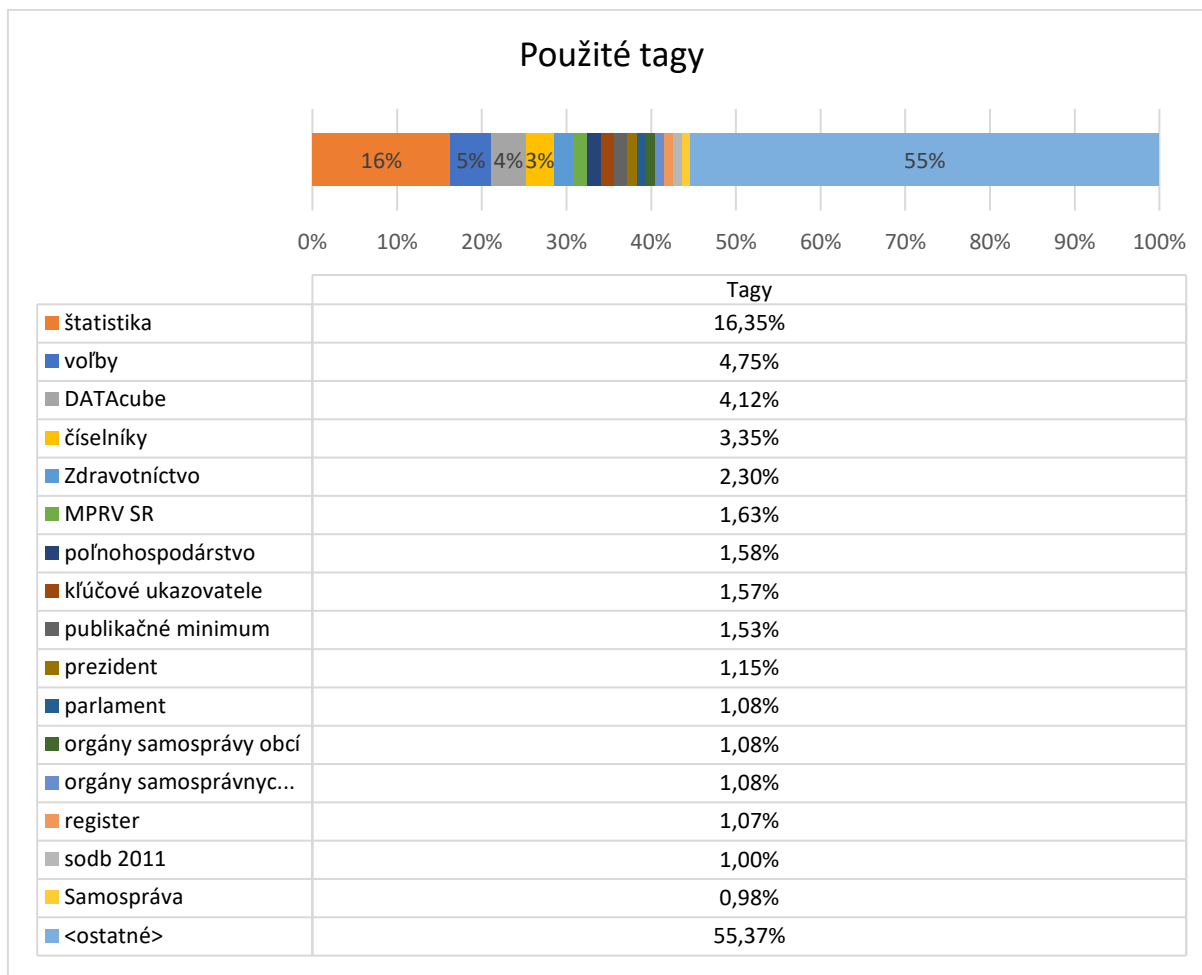
Najväčší podiel majú datasety s tagom **štatistika, voľby a DATAcube**, čo sú všetko tagy asociované najmä s datasetmi Štatistického úradu SR. V populárnych tagoch sa tiež vyskytuje **zdravotníctvo, poľnohospodárstvo, publikačné minimum, parlament či prezident**.

Celkový počet tagov je až 1088, v tagoch sa vyskytujú duplicity či preklepy. Je otázne, či systém v takomto šate plní účel jednoduchého hľadania podľa kľúčových slov. Prínosné by bolo tagy prečistiť a zmenšiť ich počet alebo motivovať OVM, aby si vybrali prioritne z preddefinovaného

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

zoznamu tagov (na tento účel by sa mohli využiť rôzne číselníky, napríklad zoznam odvetví a SK NACE alebo zoznam populárnych tagov).

Graf 4



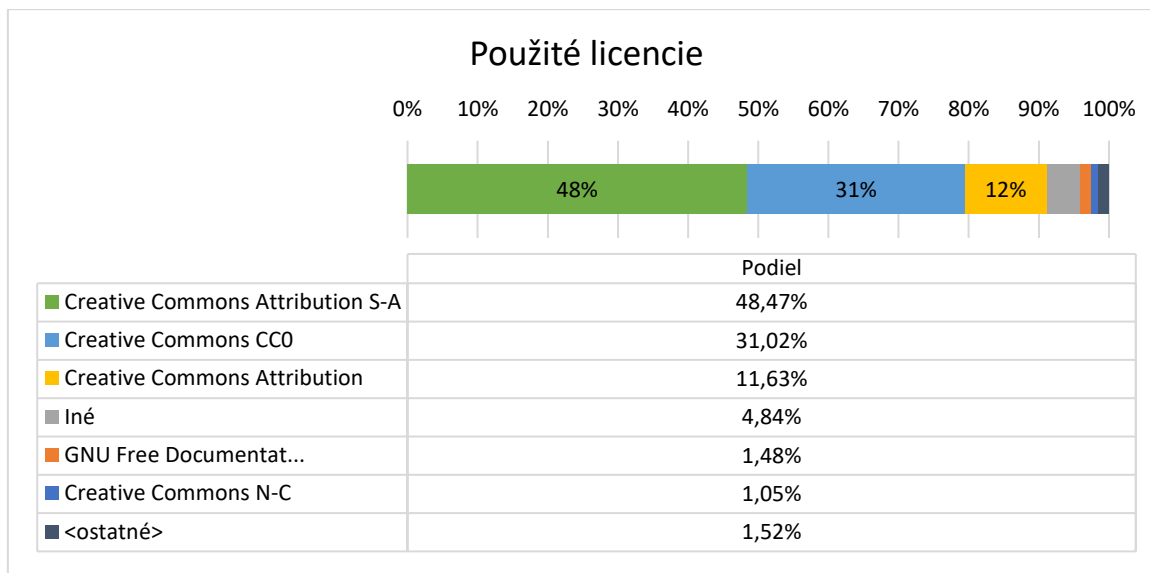
V porovnaní s analýzou za minulý rok neboli zaznamenané zmeny v najpopulárnejších tagoch. **Vznikli niektoré nové tagy napr. DATAcube alebo MPRV SR**, ktoré sa zaradili medzi najpopulárnejšie. Celkový počet jednotlivých tagov vzrástol o 97 % z 3047 na 6001, čo znamená že funkcia tagovania sa aktívne využíva.

2.4 Licencie datasetov

Licencia znamená právny režim použitia údajov. Aby používateľ vedel, ako môže údaje používať a čo musí dodržať, je nutné aby boli pri datasetoch uvedené licencie. Licencia by mala umožňovať komerčné aj nekomerčné použitie údajov, čo spĺňajú napríklad licencie Creative Commons CC0 alebo CC-BY.

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 5

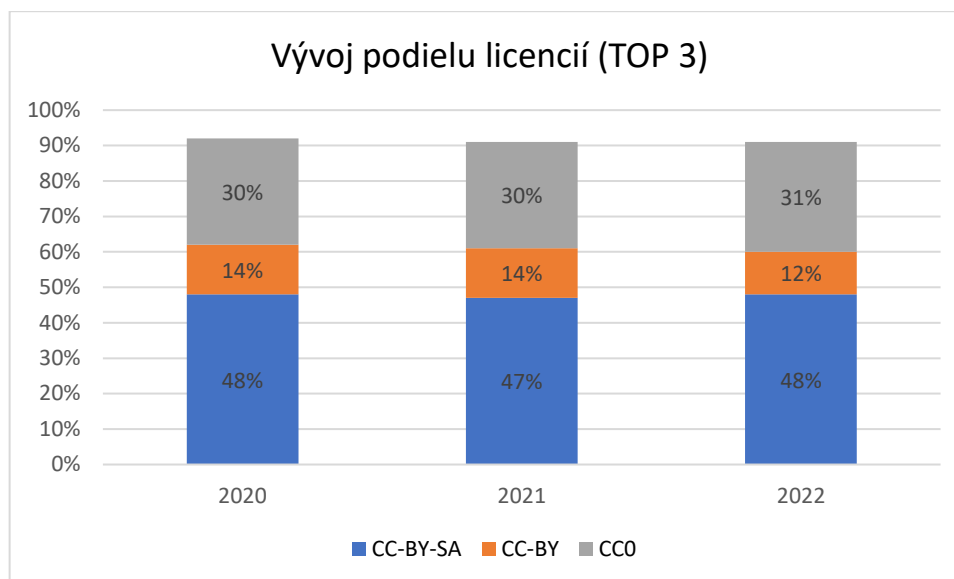


Najpoužívanějšía licencia je na portáli **CC-BY-SA**, nasleduje **CC0** a tretia je **CC-BY**. Nezanedbateľná časť datasetov používa inú licenciu ako GNU a CC, ide o 6,25 %. Pri 0,11 % datasetov nie je licencia uvedená.

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. Pomerne veľká časť datasetov používa neodporúčané licencie, je to až 58 %. Úrad vlády, ktorý bol spomínaný v minulej správe ako úrad, ktorý neuvádza licencie, momentálne používa licenciu CC pri 19 z 22 datasetov, pri 3 datasetoch neuvádza licenciu.

⁹Vyhláška 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>

Graf 6



Podiel licencií sa príliš nemení, viditeľné sú len drobné zmeny, kedy rástol podiel CC0, znižoval sa podiel CC-BY a stagnoval podiel CC BY-SA. Tento vývoj nie je pozitívny, žiaduce by bolo, aby sa licencie rýchlejšie preklápali smerom k odporúčaným licenciám CC-0 a CC-BY.

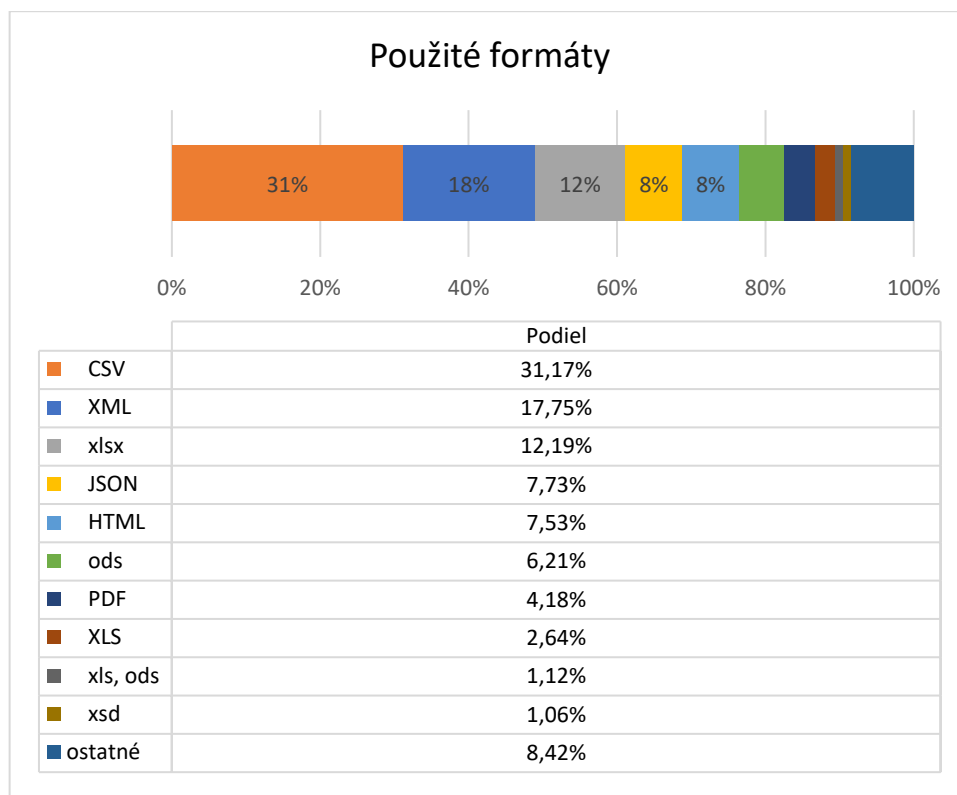
2.5 Formáty datasetov

Formát je spôsob uloženia dát (štruktúra) do súboru. Formát čiastočne určuje strojovú spracovateľnosť údajov, odporúčajú sa formáty, ktoré sú otvorené a strojovo spracovateľné (CSV, XML, JSON).

Dôležité je poznamenať, že počet datasetov sa nezhoduje s počtom súborov (jeden dataset môže obsahovať viac súborov/dátových zdrojov).

Portál data.gov.sk v záložke Datasets umožňuje filtrovať datasety podľa formátu súboru. 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.

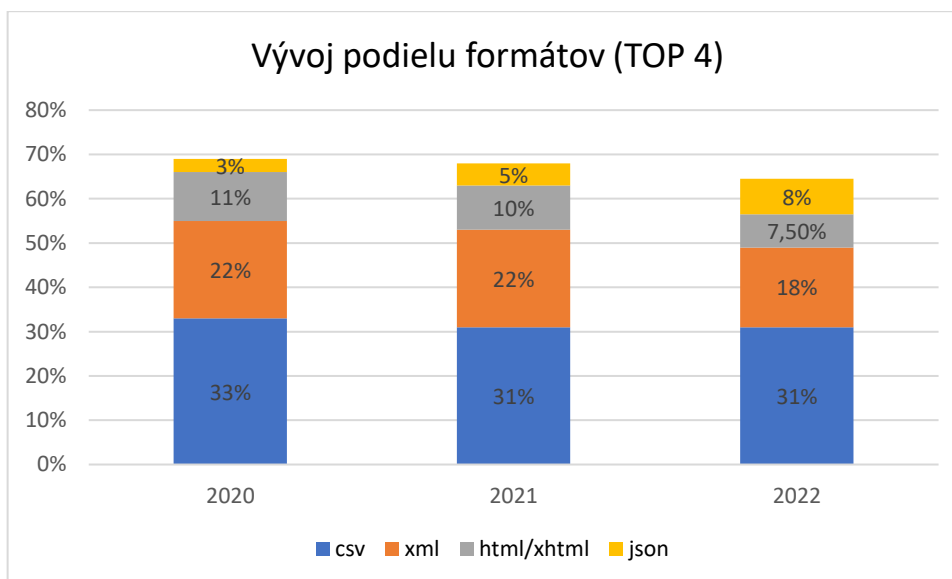
Graf 7



Najpopulárnejšie formáty sú CSV, XML, XLSX a JSON, ktoré spolu tvoria 2/3 všetkých datasetov. **Stovky datasetov sú ešte stále zverejňované v strojovo veľmi zle čitateľných formátoch PDF, HTML, DOCX.** Okrem týchto najpoužívanějších formátov sa používajú aj mnohé iné formáty, ktoré sú strojovo spracovateľné, napr. geojson, kml, shapefile a ďalšie.

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“. Taktiež by sa ako formát nemalo používať označenie zip alebo url, http, keďže zip súbor slúži len ako schránka, ktorá obsahuje iné typy súborov. Označenie http či url smeruje na webovú adresu, kde sú takisto uložené súbory iných formátov alebo je k dispozícii API.

Graf 8



Od roku 2020 viac ako dvojnásobne narástol podiel formátu JSON, čo hodnotíme kladne, keďže ide o formát ktorý je dobre strojovo spracovateľný. Pokles zaznamenal strojovo zle spracovateľný formát html/xhtml a taktiež mierny pokles formát XML.

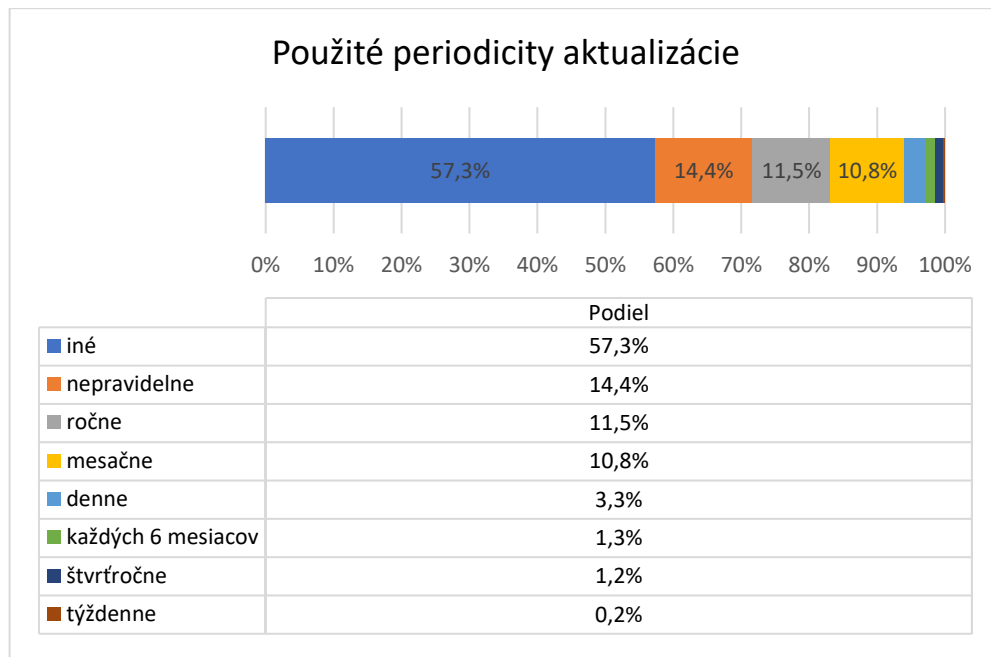
2.6 Aktualizácia datasetov

Aktualizácia datasetu je metaúdaj, ktorý hovorí o tom, ako často je dataset aktualizovaný. Na portáli je hodnota vyjadrená slovom, napríklad týždenne alebo mesačne. Prirodzene, čím častejšie je dataset aktualizovaný tým lepšie, nie všetky datasety ale treba aktualizovať rovnako často (napr. dataset s menami ministrov nie je nutné aktualizovať denne, ale postačuje napr. týždenne alebo mesačne). **Aktualizácia datasetu spolu s dátumom poslednej aktualizácie je pre používateľa jeden z najdôležitejších metaúdajov**, hovorí ako často má zmysel dataset sťahovať a či je táto periodicita dostatočná pre konkrétne použitie, ktoré má používateľ v pláne.

Na portáli má každý dátový zdroj (súbor, odkaz) vlastnú hodnotu aktualizácie, ktorú sme získali z API portálu.

Viac ako 60 % dátových zdrojov má hodnotu iné a nepravidelne, čo sú hodnoty, ktoré by sa mali využívať, ak nie je možné vybrať konkrétnejšiu periodicitu aktualizácie. Z konkrétnejších periodicít je najviac zastúpené ročne a mesačne (10-11 %), denne je aktualizovaných 3,3 % dátových zdrojov, každého pol roka a štvrťročne je to okolo 1 %, týždenne len 0,2 % dátových zdrojov.

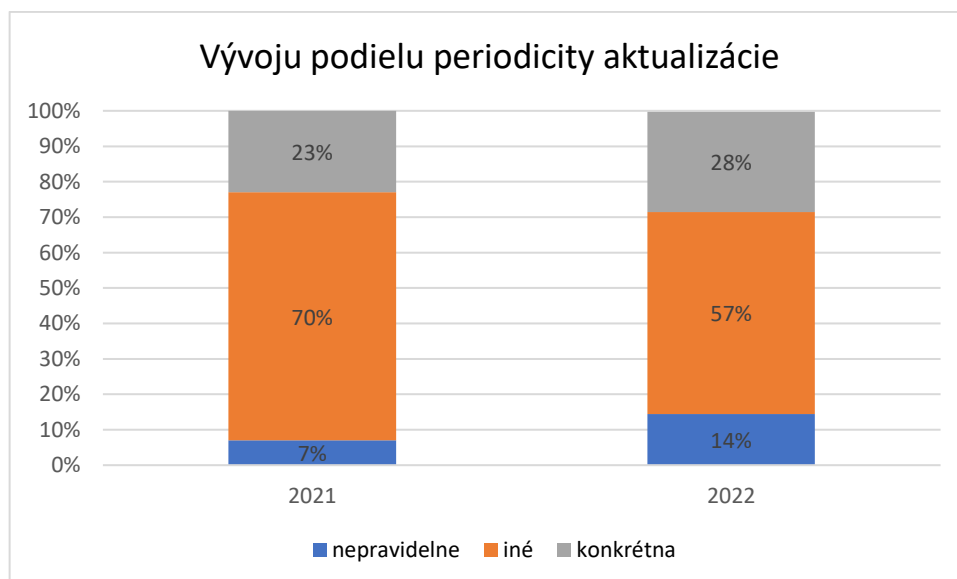
Graf 9



Zdroj: API Data.gov.sk (17.6.2022)

Vývoj podielu periodicity aktualizácie je pozitívny. **Za posledný rok narástol podiel konkrétnejších hodnôt aktualizácie z 23 % na 28 %**. Stále však ostáva väčšina datasetov ktorá má nepravidelnú aktualizáciu alebo hodnotu iné. Tieto datasety by bolo treba analyzovať a určiť, či je perióda aktualizácie určená správne, respektíve či sú datasety pravidelne aktualizované, čo je veľmi dôležité pre ich použiteľnosť.

Graf 10



2.7 Zhrnutie

Za posledný rok **pribudlo** na portáli približne **500 datasetov**, čo je pomerne veľa v porovnaní s minulými rokmi. Pozitívne je tiež to, že **2/3 väčšina OVM publikujúcich na portáli zverejnila aspoň 1 nový dataset**.

Počet datasetov však zaostáva za možnosťami, keď každý OVM by mal disponovať minimálne desiatkami, tie väčšie aj stovkami datasetov. Pomaly však pribúdajú OVM, ktoré sa k týmto číslam blížia, je to najmä Ministerstvo pôdohospodárstva, Ministerstvo zdravotníctva a Úrad pre verejné obstarávanie. Oceniť treba tiež pokrok samotného MIRRI a Protimonopolného úradu.

Zloženie licencií sa bohužiaľ príliš nemení, **stále ostáva vysoký (nadpolovičný !) podiel licencií, ktoré nie sú dostatočne otvorené. Podiel datasetov v dobre strojovo spracovateľnom formáte JSON sa stabilne zvyšuje**, z 3 % v roku 2020 na 8 % v roku 2022. Zároveň je pozitívne, že **klesá podiel formátov html/xhtml**, ktorý je zle strojovo spracovateľný z 11 % v roku 2020 na 7,5 % v roku 2022. **Stovky datasetov sú ešte stále zverejňované v strojovo veľmi zle čitateľných formátoch PDF, HTML, DOCX**. Väčšina dátových zdrojov má nastavených ako periódu aktualizácie nepravidelne alebo iné, čo môže pre používateľov znamenať problém.

3 Kvalitatívna analýza

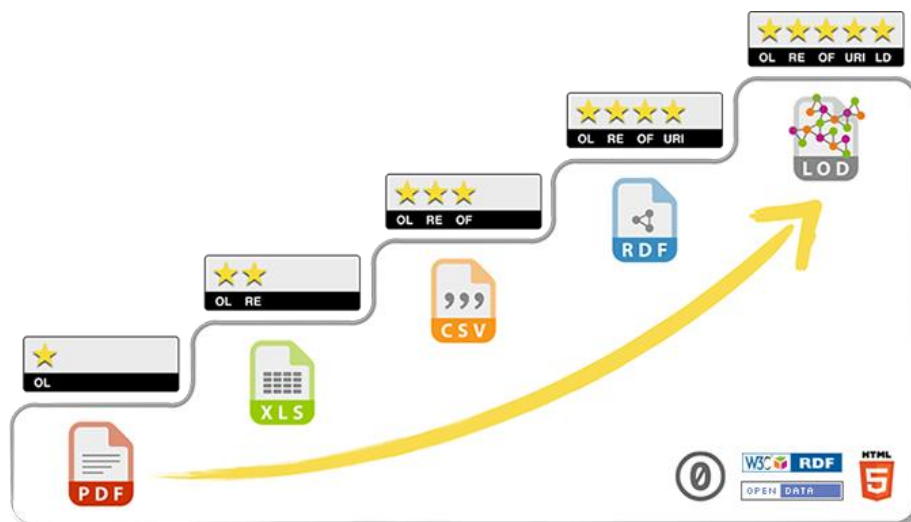
Táto kapitola analyzuje zverejňované datasety z pohľadu používateľa, podľa kvalitatívnych kritérií, ako je napr. hodnotenie podľa hviezdíčiek alebo podľa iných prekážok dátovej kvality. Zdrojom údajov a štatistík v tejto kapitole je portál data.gov.sk a minulé analýzy, pokiaľ nie je uvedené inak.

3.1 Hodnotenie datasetov podľa hviezdíčiek

Tim Berners-Lee, vynálezca webu a iniciátor prepojených dát, navrhol hodnotenie otvorených dát hviezdíčkami¹⁰:

- * 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)¹¹, § 38 až § 40.



Obrázok 2: Hodnotenie otvorených dát hviezdíčkami

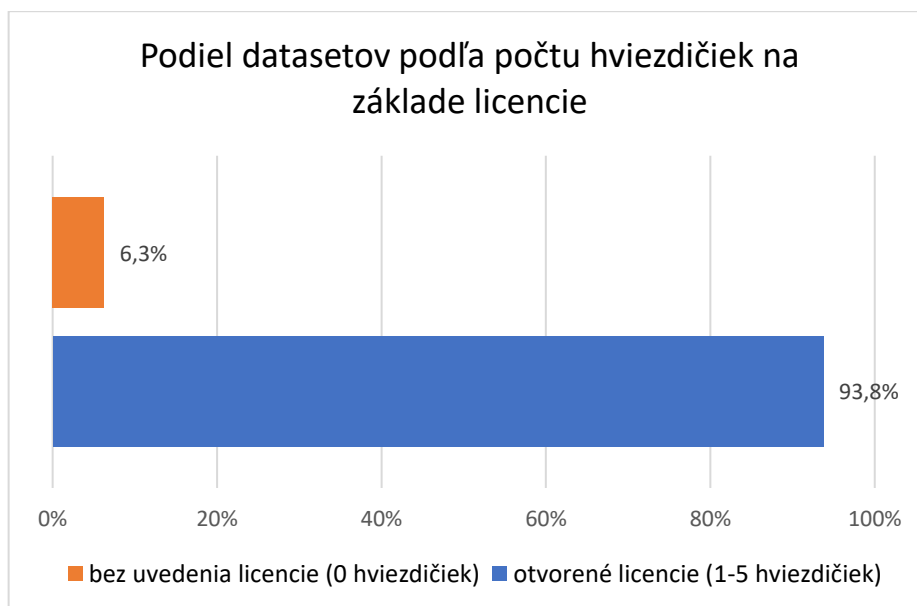
Hodnotenie hviezdíčkami možno pridať k vyššie uvedeným prehľadom licencií (kap. 2.4), formátov (kap. 2.5). Minimálna požadovaná úroveň pre otvorené dáta je úroveň 3 hviezdíčiek.

Aby bol dataset ohodnotený aspoň 1 hviezdíčkou, musí obsahovať otvorenú licenciu. Na portáli data.gov.sk sa vyskytujú licencie Creative Commons, GNU alebo Open Data Commons.

¹⁰<https://5stardata.info/cs/>

¹¹<https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2020/78/20200501>

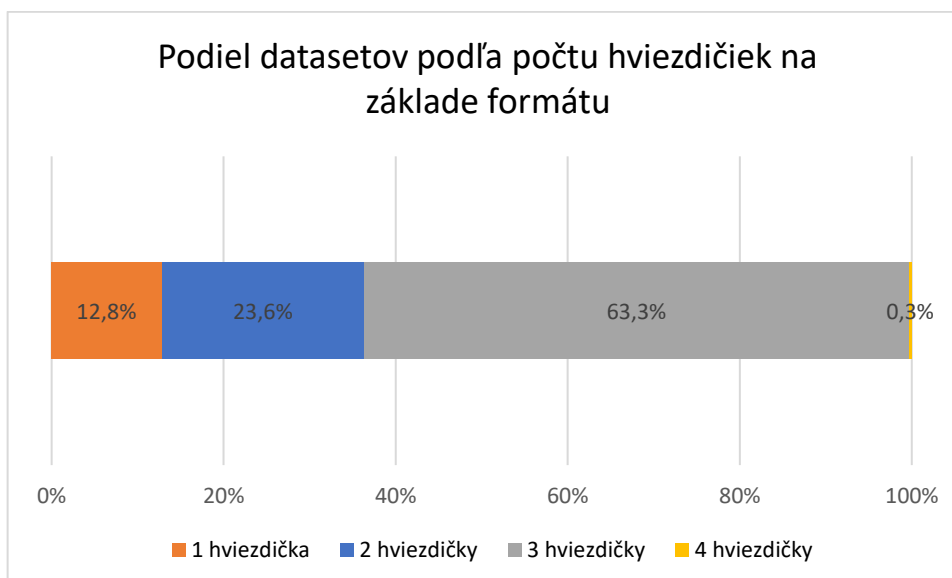
Graf 11



Oproti minulému roku **klesol podiel nevhodných (nie otvorených) licencií z 8 % na 6,3 %**. Najčastejšie je to spôsobené neuvedením licencie (uvedením hodnoty „iné“).

Ďalej je pri hodnotení hviezdíčkami dôležitý formát zverejnených datasetov hviezdíčkami. 2 hviezdíčky dostáva formát, ktorý je strojovo spracovateľný ale nie je otvorený (napr. XLSX), 3 hviezdíčky strojovo spracovateľný a otvorený formát (napr. CSV, JSON, XML), viac hviezdíčiek prepojené dáta.

Graf 12



Najväčší podiel majú 3 hviezdičky, takmer 2/3, čo sú plnohodnotné otvorené dáta. 2 hviezdičky, ktoré síce nie sú plnohodnotné otvorené dáta, ale sú ešte stále dobre strojovo spracovateľné sa vyskytujú pri takmer 24 % datasetov. Približne 13 % datasetov dosahuje len 1 hviezdičku, čo je tiež podiel datasetov, ktoré je treba nahradiť datasetmi s vyšším počtom hviezdičiek.

Formát RDF, ktorý poskytuje ontológiu pre metadáta ku datasetu, je na portáli data.gov.sk len pri 17 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 hviezdičiek je použitie ontológie pri definícii údajov (metadátach), použitie referencovateľného identifikátora (permanentnej URI/URL) pre dataset a použitie referencovateľných identifikátorov v dátach. Zverejnené 3-hviezdičkové datasety na portáli data.gov.sk majú spravidla permanentnú URI adresu, také možno považovať za 4-hviezdičkové.

Úroveň 5 hviezdičiek spĺňajú tie datasety, ktoré majú 4 hviezdičky 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. Takéto datasety zverejnilo Ministerstvo životného prostredia SR.

3.2 Najčastejšie problémy znižujúce použiteľnosť datasetov

Tu sumarizujeme najbežnejšie problémy používateľa pri používaní datasetov. Tieto problémy je možné riešiť do veľkej miery automatizovane, kontrolou metadát a aktuálnosti odkazov, čo by mal priniesť nový portál OpenData 2.0.

Sú to:

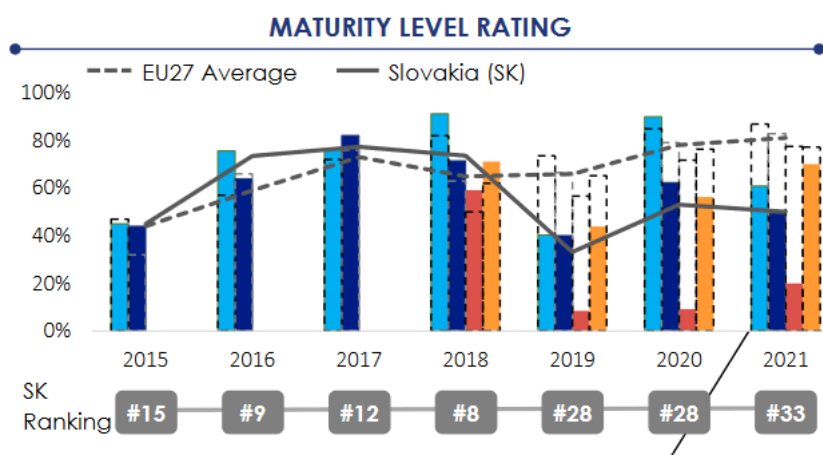
- Nedostupný dataset (chybný odkaz),
- Názov súboru sa zmenil – môže byť problém pri automatizovanom spracovaní,
- Nedostupné metadáta – metadáta o obsahu súboru, jeho vzniku a význame jednotlivých stĺpcov,
- Nedostupné metadáta – chýbajúci kontakt (e-mail, telefón), kde sa môže používateľ datasetu dotazovať. Často sa uvádza názov iba organizačnej jednotky, čo je nedostatočné.
- Neaktuálnosť dát,
- Nevhodná – neuvedená alebo nedostatočne otvorená licencia,
- Nevhodný formát, ktorý nie je dobre strojovo spracovateľný.

3.3 Rebríčky

Open data maturity index¹²

Open data maturity index je rebríček každoročne vyhodnocovaný Európskou komisiou. Slovensko sa umiestnilo v poslednom sledovanom období **2021** na **33. priečke** ako posledná z krajín EÚ, v najslabšej skupine sledovaných európskych štátov (začiatovníci) zo štyroch skupín (lídri trendu, rýchli nasledovníci trendu, nasledovníci, začiatovníci). Metodika je založená na posúdení 4 dimenzií otvorených dát: politika, portál, dopad, kvalita.

Obrázok 1



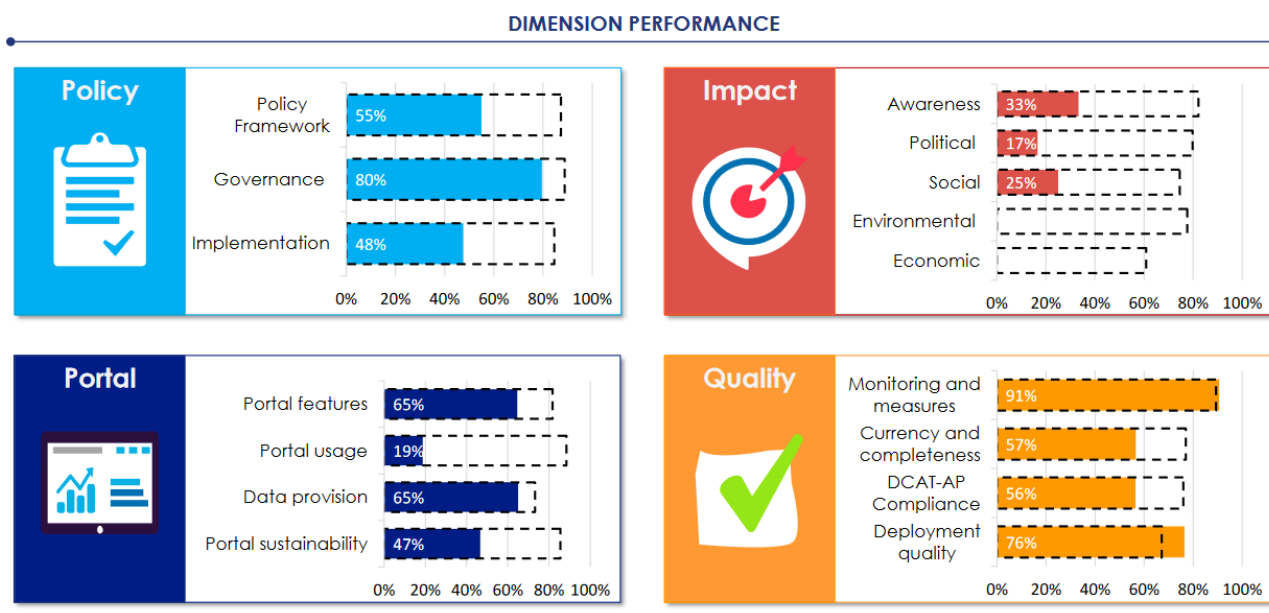
Zdroj: https://data.europa.eu/sites/default/files/country-factsheet_slovakia_2021.pdf

Slovensko sa oproti minulému roku prepadlo v hodnotení portálu aj politik, mierne si polepšilo pri dopade a kvalite.

Pri detailnejšom pohľade je zrejmé, že veľké rezervy máme pri hodnotení dopadu otvorených dát, taktiež pri hodnotení používania a udržateľnosti portálu a taktiež pri politikách, najmä pri implementácii a nastavení politik.

¹² <https://data.europa.eu/en/impact-studies/open-data-maturity>

Obrázok 2



Zdroj: https://data.europa.eu/sites/default/files/country-factsheet_slovakia_2021.pdf

Dobrá správa je, že toto hodnotenie sa môže v najbližších rokoch zlepšiť, najmä vplyvom zavedenia nového centrálného portálu otvorených dát, publikačného minima a zvýšenej pozornosti dátam s vysokou pridanou hodnotou a sledovaniu vplyvu (dopadu) otvorených dát.

OECD OURdata Index - vládne otvorené dáta ¹³

Rebríček porovnáva vládne otvorené dáta v troch oblastiach – otvorenosť, užitočnosť a znovupoužiteľnosť. Posledné dostupné dáta sú za rok 2019. Slovensko sa v tomto rebríčku umiestnilo na **27. mieste z 35 krajín**. Čo sa týka dostupnosti datasetov, podľa odporúčania autorov by sa malo Slovensko zamerať na **datasety s vysokou pridanou hodnotou**. Dostupnosť údajov by malo Slovensko podľa odporúčania zvýšiť **zlepšením portálu, najmä umožnením používateľom pridávať vlastné dáta a vytvorením diskusného fóra**. V oblasti podpory znovupoužívania údajov by sme sa mali zamerať na **vzdelávanie štátnych zamestnancov**.¹⁴

Open data watch

Slovensko sa umiestnilo v poslednom sledovanom období **2020/21¹⁵ na 10. mieste** so 78 % v rebríčku 187 hodnotených krajín (Singapur na 1. mieste s 92 %). Lepšie ako Slovensko dopadlo Poľsko, horšie Maďarsko, Rakúsko aj Česko. Oproti predošlému obdobiu 2018 sa Slovensko zlepšilo. Posudzujú sa oficiálne zverejňované štatistiky štátu v oblastiach: sociálna, ekonomická, životné prostredie. Prehľad o Slovensku za rok 2020¹⁶ je na základe dát Štatistického úradu SR¹⁷. **Najhoršie**

¹³ <https://www.oecd.org/gov/digital-government/open-government-data.htm>

¹⁴ <https://www.oecd.org/gov/digital-government/ourdata-index-slovak-republic.pdf>

¹⁵ <https://odin.opendatawatch.com/>

¹⁶ <https://odin.opendatawatch.com/Report/Reports>

¹⁷ <http://www.statistics.sk>

hodnotenie zaznamenalo Slovensko v oblasti Bezpečnosť potravín a výživa, Výsledky vzdelávania, Výsledky zdravotníctva a Výstavba.

Index digitálnej ekonomiky a spoločnosti (DESI)

V indexe digitálnej ekonomiky a spoločnosti (DESI) 2021¹⁸ patrí Slovensku **22. miesto v rebríčku 27 krajín EÚ, rovnaké ako pred rokom. Slovensko¹⁹ je v hodnotení DESI pod priemerom EÚ s podpriemerným rastom za roky 2015 – 2021, 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 Dánsko, Fínsko a Švédsko.

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 máji 2021 vláda schválila stratégiu a akčný plán na zlepšenie postavenia Slovenska v indexe DESI do roku 2025²⁰. Akčný plán obsahuje tieto úlohy súvisiace s otvorenými dátami:

- V.1 Implementácia národného projektu “Otvorené údaje 2.0 – Rozvoj centrálnych komponentov pre kvalitné zabezpečenie otvorených údajov“,
- V.2 Implementácia národného projektu „Rozvoj platformy integrácie údajov (centrálne integračná platforma) a Manažment osobných údajov“,
- V.4 Implementácia národného projektu „Dátová integrácia: sprístupnenie údajovej základne VS vrátane otvorených údajov prostredníctvom platformy dátovej integrácie

Ostatné rebríčky

V minulosti boli v tejto analýze vyhodnocované aj rebríčky Open data barometer²¹ a Global open data index²². Posledné hodnotenie Slovenska je však v nich staré viac ako 5 rokov, preto ich pre neaktuálnosť v tohtoročnej analýze nezaraďujeme.

3.4 Zhrnutie

Väčšina datasetov na portáli v súčasnosti z hľadiska formátu dosahuje 3 hviezdičky. **Až 12 % datasetov má však kvalitu len 1 hviezdičky a 23,5 % datasetov len 2. Najmä datasety s 1 hviezdičkou je potrebné prioritne zverejniť s vyšším počtom hviezdičiek.** Počet datasetov

¹⁸<https://ec.europa.eu/newsroom/dae/redirection/document/80563>

¹⁹ <https://ec.europa.eu/newsroom/dae/redirection/document/80599>

²⁰ <https://www.mirri.gov.sk/sekcie/informatizacia/digitalna-transformacia/strategia-a-akcny-plan-na-zlepsenie-postavenia-slovenska-v-indexe-desi-do-roku-2025/index.html>

²¹ <https://opendatabarometer.org>

²² <https://index.okfn.org/>

s viacerými hviezdčkami stúpa len pomaly, zásluhou najmä MIRRI. V súčasnosti sa ich podiel počíta v desatinách percenta.

Väčšina datasetov má uvedenú licenciu, **existuje však 6,3 % datasetov kde licencia nie je uvedená alebo je uvedené iná. Tieto datasety je tiež potrebné prioritne zverejniť s jasne uvedenou licenciou (najlepšie čo najotvorenejšou).**

Posun v sledovaní kvality údajov sa očakáva od nového portálu, ktorý je v príprave. Tento portál by mal mať schopnosť automaticky sledovať a vyhodnocovať kvalitu na základe metaúdajov, upozorňovať na nefunkčné odkazy apod. (problémy pomenované v kapitole 3.2.).

Slovensko sa prepadlo v rebríčku Open data maturity na poslednú priečku a nemá tak už kam klesať. **V záujme zlepšenia postavenia v rebríčku je potrebné iniciovať najmä posilnenie politík smerujúcich k implementácii otvorených dát, sledovanie dopadu otvorených dát a taktiež lepšie monitorovať a udržiavať portál.** Riadiť by sme sa mali aj odporúčaniami OECD²³, ktoré hoci sú 3 roky staré, sú stále relevantné.

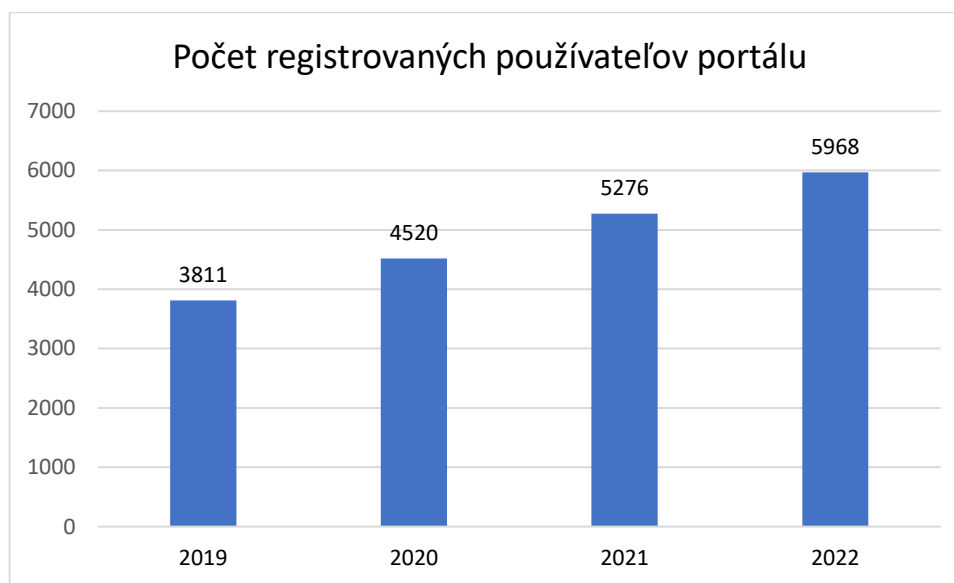
²³ <https://www.oecd.org/gov/digital-government/open-government-data.htm>

4 Používanie portálu

4.1 Počet registrovaných používateľov portálu

Na portáli data.gov.sk v záložke Štatistiky sa uvádza počet registrovaných používateľov portálu. K 24.6.2022 bol tento počet **5 968**, čo je o **692 (13 %)** viac ako pred rokom. Počet registrovaných používateľov v posledných 4 rokoch rastie rovnomerne.

Graf 13



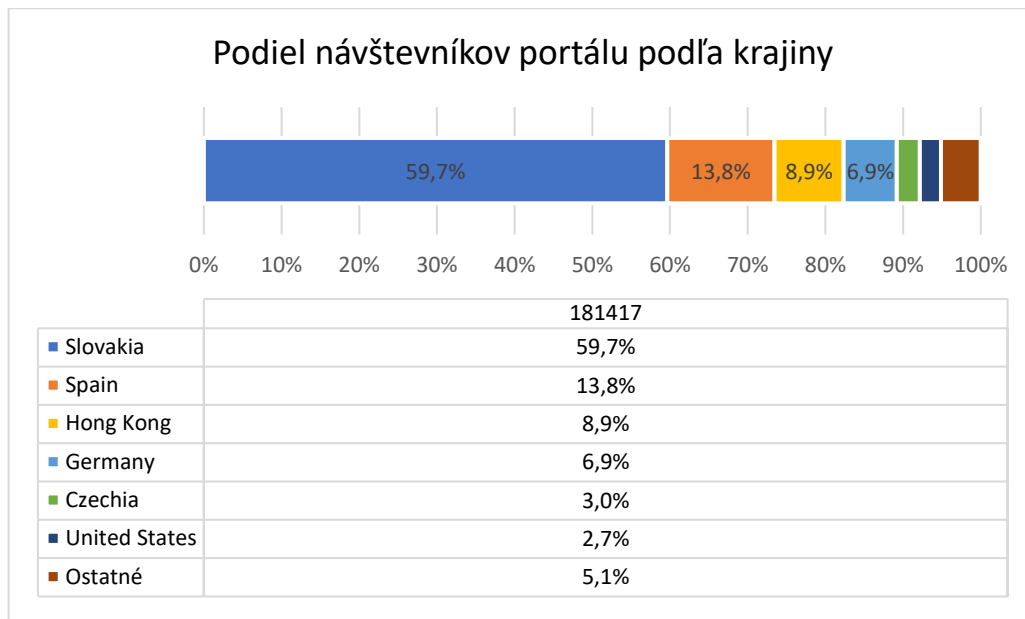
Zdroj: Data.gov.sk a minulé analýzy

4.2 Návštevníci podľa krajín

Počet návštevníkov na portáli v období od 1.1.2017 do 15.6.2022 bol 586 000. Z týchto návštevníkov sme vylúčili návštevníkov z Indonézie a Nepálu, kde predpokladáme že sa nejednalo o skutočných ľudí (priemerné trvanie návštevy bolo 0,01 s.). Brali sme do úvahy zvyšok, **304 132 návštevníkov**.

Najviac návštevníkov pochádzalo zo Slovenska – takmer 60 %, 181 000. Druhou krajinou s najvyšším počtom návštevníkov bolo Španielsko s podielom takmer 14 %. Na treťom mieste bol zaznamenaný Hong-Kong s podielom takmer 9 %. Ďalšie najčastejšie krajiny sú Nemecko (6,9 %), Česko (3 %) a Spojené štáty Americké (2,7 %). Na ďalších miestach s viac ako tisícovkou návštevníkov sú Čína, Singapur, Spojené kráľovstvo a Rakúsko.

Graf 14



Zdroj: NASES

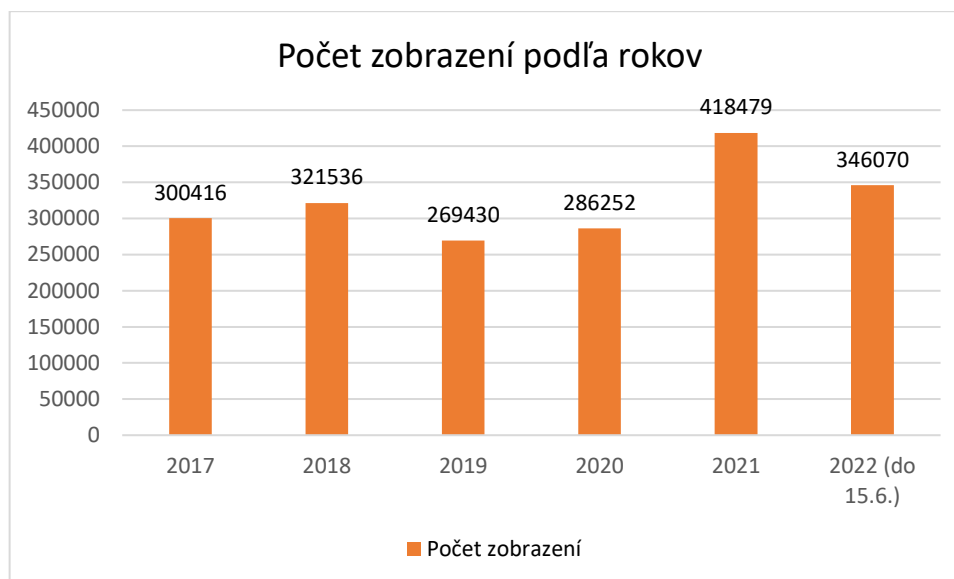
4.3 Počet zobrazení podľa rokov

Počet zobrazení stránok meria, koľko stránok bolo na portáli načítaných. Jeden používateľ mohol zobraziť jednu alebo aj viac stránok. Nedá sa teda porovnávať s počtom používateľov, lebo zachytáva aj počet používateľov aj mieru ich aktivity.

Počet zobrazení stránok je k dispozícii od 1.1.2017 do 1.6.2022. Pokiaľ sa pozrieme na čísla vyjadrené v grafe po rokoch, vidíme, že počet zobrazení v rokoch 2017 – 2020 stagnoval, ba až klesal. To koreluje s počtom datasetov, ktorý v týchto rokoch tiež stagnoval. **V rokoch 2021 a 2022 počet datasetov rastie, rastie aj počet zobrazení, v roku 2021 až o 44 % v porovnaní s predchádzajúcim rokom.** Rok 2022 zatiaľ je len približne v polovici, v prípade ak by tempo rastu zodpovedalo prvému polroku, môžeme hovoriť o 600-700 tis. zobrazeniach za celý rok 2022, čo je nárast oproti roku 2021 o viac ako 50 %. Tieto čísla však musíme interpretovať opatrne, nevieme presne či sa v nich nenachádzajú aj zobrazenia z krajín popísaných v kapitole 4.2., kde je predpoklad že nejde o reálnych ľudí.

Počet zobrazení sme skúmali aj podľa jednotlivých mesiacov v roku, nezistili sme však žiadne opakujúce sa vzorce. V niektorých rokoch bol počet zobrazení vyšší v posledných mesiacoch v roku, inokedy to boli jarné alebo letné mesiace.

Graf 15



Zdroj: NASES

4.4 Najpopulárnejšie datasety

Cieľom portálu otvorených dát je umožniť čo najširšiu použiteľnosť dát rôznymi skupinami používateľov. Jedným z najlepších indikátorov toho, ako sú dáta používané je počet ich stiahnutí²⁴. Tento údaj nemáme bohužiaľ od prevádzkovateľa portálu (NASES) k dispozícii, máme však k dispozícii **počet zobrazení v roku 2021**. To síce nemusí znamenať že používateľ dataset aj naozaj stiahol, ale znamená to že sa o neho minimálne zaujímal.

Z hľadiska najpopulárnejších datasetov sa na prvých priečkach vyskytujú Register adries, vestník Úradu priemyselného vlastníctva SR, zoznam liekov, počet prístupov UPVS za deň, zoznam vyplatených faktúr Pôdohospodárskej platobnej agentúry. Populárne sú aj datasety Národnej rady Slovenskej Republiky – programy schôdzí, zoznam rozhodnutí Protimonopolného úradu, register budov alebo register poľovných revírov.

Tabuľka 2: Počet zobrazení datasetov na portáli (TOP 20)

P.	Názov datasetu	Vlastník	Počet zobrazení
1	Register adries - Zmenové dávky registra adries	Ministerstvo vnútra SR	35216
2	Vestník Úradu priemyselného vlastníctva SR	Úrad priemyselného vlastníctva Slovenskej republiky	21058
3	Hlavný zoznam liekov s platnou registráciou na území SR a 6 ďalších prehľadných zoznamov (ATC, FIRMA, IS, ŠTÁTY, TYPY REGISTRÁCIÍ a VÝDAJE)	Štátny Ústav pre kontrolu liečiv	5806

²⁴ Nejde o jediný indikátor, dôležité sú aj konkrétne použitia dát, ich prínos, počet používateľov a ich príslušnosť k používateľským skupinám, vytvorená pridaná hodnota, atď. Všetky tieto dodatočné kritériá sa však oveľa ťažšie zisťujú.

4	Počet prístupov na ÚPVS za deň	Národná agentúra pre sieťové a elektronické služby - NASES	3354
5	Register adries - Zmenové dávky	Ministerstvo vnútra SR	3089
6	Počet prístupov na ÚPVS	Národná agentúra pre sieťové a elektronické služby – NASES	2681
7	Zoznam vyplatených faktúr PPA	Pôdohospodárska platobná agentúra	2485
8	Register Adries - Register ulíc	Ministerstvo vnútra SR	2321
9	Register Adries - Register obcí	Ministerstvo vnútra SR	2026
10	Register Adries - Register vchodov (orientačných čísiel)	Ministerstvo vnútra SR	1893
11	Programy schôdzí 8. volebného obdobia	Kancelária Národnej rady SR	1417
12	Programy schôdzí 7. volebného obdobia	Kancelária Národnej rady SR	1365
13	Zoznam rozhodnutí PMÚ SR	Protimonopolný úrad SR, Bratislava	1224
14	Register Adries - Register budov (súpisných čísiel)	Ministerstvo vnútra SR	1176
15	Register poľovných revírov (PR)	Národné lesnícke centrum	1101
16	Adresy podľa krajov	Ministerstvo vnútra SR	1026
17	Zoznam faktúr PMÚ SR	Protimonopolný úrad SR, Bratislava	998
18	Zoznam OVM s pridelenou rolou na vytváranie listinného rovnopisu v schránke	Národná agentúra pre sieťové a elektronické služby – NASES	998
19	Zoznam objednávok PMÚ SR	Protimonopolný úrad SR, Bratislava	938
20	Vestník verejného obstarávania marec 2016	Úrad pre verejné obstarávanie	937

Zdroj: NASES

Spomedzi všetkých datasetov zverejnených OVM sme vybrali ten najpopulárnejší podľa najvyššieho počtu zobrazení. Najpopulárnejšie datasety zverejňujú Kancelária NR SR, MIRRI, Ministerstvo pôdohospodárstva a rozvoja vidieka, Ministerstvo vnútra SR, Ministerstvo zdravotníctva SR a Úrad priemyselného vlastníctva SR.

Tabuľka zachytáva rok 2021, takže OVM, ktoré začali výraznejšie zverejňovať otvorené dáta neskôr, v nej nie sú zachytené. OVM, ktoré napriek tomu dosiahli nižšie čísla, by sa mali zamyslieť nad tým, či nemajú vo svojej gescii dáta, ktoré je možné zverejniť ako otvorené a ktoré by mohli používateľov zaujímať.

Tabuľka 3: Najpopulárnejšie datasety OVM

P.	Vlastník	Najpopulárnejší dataset	Počet zobrazení
1	Ministerstvo vnútra SR	Register adries - Zmenové dávky registra adries	35216
2	Úrad priemyselného vlastníctva Slovenskej republiky	Vestník Úradu priemyselného vlastníctva SR	21058
3	Ministerstvo zdravotníctva SR	Hlavný zoznam liekov s platnou registráciou na území SR a 6 ďalších prehľadných zoznamov (ATC, FIRMA, IS, ŠTÁTY, TYPY REGISTRÁCIÍ a VÝDAJE)	5806
4	Ministerstvo investícií, regionálneho rozvoja a informatizácie Slovenskej republiky	Počet prístupov na ÚPVS za deň	3354
5	Ministerstvo pôdohospodárstva a rozvoja vidieka SR	Zoznam vyplatených faktúr PPA	2485
6	Kancelária Národnej rady SR	Programy schôdzí 8. volebného obdobia	1417
7	Protimonopolný úrad SR	Zoznam rozhodnutí PMÚ SR	1224

8	Úrad pre verejné obstarávanie	Vestník verejného obstarávania marec 2016	937
9	Ministerstvo dopravy a výstavby SR	Zoznam schválených ŽoNFP OPII	883
10	Ministerstvo školstva, vedy, výskumu a športu Slovenskej republiky	Centrálny register študentov	515
11	Štatistický úrad SR	Obce Slovenskej republiky	379
12	Ministerstvo spravodlivosti SR	Priemerná dĺžka konania	297
13	Ministerstvo kultúry Slovenskej republiky	Register múzeí a galérií Slovenskej republiky	276
14	Ministerstvo obrany SR	Faktúry MO SR	275
15	Úrad geodézie, kartografie a katastra SR	Zoznam Datasetov ÚGKK SR	218
16	Ministerstvo životného prostredia SR	CITES - výročné správy	216
17	Všeobecná zdravotná poisťovňa	VSŽP - zoznamy dlžníkov	202
18	Ministerstvo financií SR	041 Dotácie poskytnuté z kapitoly VPS	169
19	Ministerstvo zahraničných vecí a európskych záležitostí SR	Elektronické konzulárne služby	168
20	Úrad vlády SR	Rozpočtová rezerva predsedu vlády SR	152
21	Ministerstvo práce, sociálnych vecí a rodiny SR	Schválené národné projekty	116
22	Úrad pre normalizáciu, metrológiu a skúšobníctvo Slovenskej republiky	Zoznam všetkých datasetov ÚNMS SR	87
23	Ministerstvo hospodárstva SR	Zoznam MH SR využívaného otvoreného softvéru	55
24	Správa štátnych hmotných rezerv Slovenskej republiky	Zoznam datasetov SŠHR SR	51
25	Úrad jadrového dozoru Slovenskej republiky	Objednávky	44
26	Národný bezpečnostný úrad	Zoznam datasetov NBÚ	23
27	Generálna prokuratúra SR	Zoznam odsúdených právnických osôb	10
28	Najvyšší kontrolný úrad SR	NKÚ SR Zoznam objednávok	4

Zdroj: NASES

4.5 Počet aplikácií

Portál umožňuje registráciu aplikácií²⁵ vytvorených s použitím otvorených dát. Bohužiaľ na zaregistrovanie aplikácie je **nutné prihlásenie občianskym preukazom s elektronickým čipom**, čo je bariéra ktorá čiastočne môže za nízky počet zaregistrovaných aplikácií. V minulom roku bolo aplikácií len 5, v tomto roku je ich aktuálny počet 11, čo je tiež veľmi málo. Aplikácií nad otvorenými dátami existuje nepochybne viac, zoznamy je možné nájsť na stránkach neziskových organizácií²⁶. Taktiež samosprávy vytvárajú vlastné aplikácie²⁷.

4.6 Zhrnutie

Z hľadiska návštevnosti v posledných 2 rokoch portál zaznamenal výrazné zvýšenie záujmu. Okrem problému chýbajúcich alebo ťažko dostupných funkcií (po prihlásení cez el. OP), **portál publikuje málo dát o svojej činnosti.** OVM a dátoví kurátori nevedia, koľko ľudí používa ich datasety, používatelia dát nevedia, ktoré dáta sú aktuálne žiadané a nevedia ani zistiť, ktoré

²⁵ <https://data.gov.sk/apps>

²⁶ <https://utopia.sk/wiki/display/opendata/SK+a+CZ+opendata+aplikacie+a+projekty>, <https://www.alvaria.sk/open-data-aplikacie/>

²⁷ <https://geopresovregion.sk/home/>, <https://opendata.kosice.sk/pages/vizualizacie>

Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2022

datasety sú na portáli populárne. Chýbajú dodatočné funkcie, ktoré majú spríjemniť používanie a vtiahnuť do používania portálu aj používateľov, ktorí nehľadajú presný dataset alebo im nie je téma otvorených dát dôverne známa.

Ako najpopulárnejšie datasety za rok 2021 figurujú Register Adries, Vestník Úradu priemyselného vlastníctva a Zoznam liekov.

5 Štandardy a iniciatívy v oblasti otvorených dát

V tejto časti mapujeme dôležité strategické a legislatívne dokumenty, štandardy a verejné či mimovládne iniciatívy v oblasti zverejňovania a využívania otvorených dát.

5.1 Publikačné minimum štátnej správy

Publikačné minimum (PM) slúži ako štandard pre OVM, podľa ktorého sa vie orientovať, ktoré datasety je vhodné zverejňovať²⁸. Ďalším nemenej dôležitým benefitom publikačného minima, ak bude definované aj na úrovni štruktúry datasetov je vyššia kompatibilita datasetov rozličných OVM, jednoduchšia práca s dátami a ich spájanie.

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 je vytvorené, zverejnené²⁹ a časť datasetov sa z neho publikuje v súlade s úlohou 5. Publikačné minimum obsahuje 37 datasetov v 4 oblastiach: Vláda a verejný sektor, Ekonomika a financie, Spravodlivosť a Obyvateľstvo a spoločnosť. Časť datasetov (12) je povinných od 1.1.2022, časť bude povinná až prijatím zákona o údajoch, čo bude pravdepodobne najskôr na jeseň 2022.

Úroveň zverejňovania datasetov PM je rôzna, napríklad pri organizačnej štruktúre je vyššia, kým napríklad pri zahraničných pracovných cestách je nízka – dáta zverejňuje len 1 OVM. **Priemerne sa však pohybuje okolo hranice 9 OVM, čo je približne 32 % OVM. Túto aktivitu hodnotíme kladne, keďže sa podarilo významnú časť OVM motivovať, aby začali zverejňovať ďalšie datasety.** Veľkou otázkou, ktorou by sa malo zaoberať MIRRI, zostáva kvalita týchto datasetov, ich použiteľnosť a aké aplikácie alebo vizualizácie či využitia by mohli resp. budú nad nimi vznikajúť.

Tabuľka 4: Počet OVM publikujúcich datasety publikačného minima štátnej správy

Dataset	Počet OVM
Objednávky	17
Faktúry	16
Dotácie (schémy, obmedzenia, výzvy, žiadosti)	14

²⁸ OVM môže a má zverejňovať aj datasety mimo Publikačného minima, minimum slúži ako základná báza

²⁹ <https://metais.vicemier.gov.sk/publicspace?pageId=67152325>

Organizačná štruktúra	12
Úradné hodiny	11
Strategické dokumenty	9
Pracovné skupiny	8
Podriadené organizácie	7
Kontaktné miesta	6
Najžiadanejšie otvorené údaje OVM	6
Prijaté žiadosti podľa Infozákona	5
Zahraničné služobné cesty	1
Zoznam všeobecne záväzných právnych predpisov	0

Zdroj: Data.gov.sk

5.2 Datasets s vysokou pridanou hodnotou

Smernica EP č. 2019/1024³⁰ 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**. Tieto datasety majú byť na základe smernice zverejňované prednostne.

Ide o jeden z typov datasetov, ktoré majú vysoký potenciál pre znovupoužitie. V súčasnosti portál neumožňuje rozlíšiť tieto datasety od ostatných napríklad na základe filtrovania, nevieme teda určiť ich presný počet, prípadne ktoré OVM ich zverejňujú. Možné by to bolo napríklad, ak by OVM všetky takéto datasety otagovali dohodnutým tagom alebo ak by to technicky umožnil portál.

5.3 Prieskum dopytu verejnosti po najžiadanejších datasetoch a API

Úrad splnomocnenca vlády SR pre rozvoj občianskej spoločnosti na základe úloh v akčných plánoch Iniciatívy pre otvorené vládnutie v SR každoročne od roku 2015 vykonáva prieskum dopytu verejnosti po najžiadanejších datasetoch a API.³¹ Najžiadanejšie datasety a API by mali byť zverejňované prioritne, keďže sa predpokladá ich vysoká využiteľnosť.

Podľa vyhodnotenia získaného od MIRRI bolo zverejnených 21 datasetov a API, 71 datasetov a API nebolo zverejnených, čo je zverejnenie na úrovni 22,8 %.

Pri prieskumoch bol doteraz problémom aj fakt, že mnohí respondenti nedefinovali svoje požiadavky dostatočne precízne, čím vznikali časté pochybnosti, čo presne by malo byť zverejnené. Prieskumy by preto do budúcnosti mali obsahovať aj kontakt na respondenta, aby OVM, ÚSV ROS či Dátová kancelária mohli priamo komunikovať s respondentom/mi a zistiť, o aké dáta má záujem.

³⁰https://datalab.digital/wp-content/uploads/Smernica2019_1024OPENDATA.pdf

³¹<https://datalab.digital/prieskum-najzidanejsich-datasetov/>

Ďalším problémom prieskumov je nízka motivácia OVM zverejňovať prioritné datasety a API, ktoré z prieskumu vyplynú. Mnohé požiadavky na zverejnenie dát sa počas rokov opakovali a napriek tomu neboli doteraz zverejnené. Príkladom je dopravná nehodovosť, ktorá sa v prieskume vyskytla prvýkrát v roku 2015³², napriek tomu tieto dáta neboli doteraz zverejnené vo formáte otvorených dát.

Nový Akčný plán Iniciatívy pre otvorené vládnutie na roky 2022-2024 (v čase písania tejto analýzy v MPK) obsahuje úlohu B.4. pre MIRRI SR „v spolupráci so splnomocnencom vlády SR pre rozvoj občianskej spoločnosti zmapovať a vyhodnotiť kvalitatívne a kvantitatívne dáta vo vybraných oblastiach na základe vypracovanej metodiky a **zmapovať a vyhodnotiť dopyt verejnosti po najžiadanejších otvorených dátach verejnej správy**“ s termínom plnenia 31.3.2023.

5.4 Národná koncepcia informatizácie verejnej správy

Národná koncepcia informatizácie verejnej správy (NKIVS), resp. jej aktualizácia bola schválená na rokovaní vlády [uznesením vlády SR č. 763/2021](#) zo 14.12.2021. Dokument definuje hlavné problémy, výzvy a priority informatizácie. V oblasti otvorených údajov (strategická priorita Manažment údajov) je definovaný princíp „čo nie je tajné je verejné“, dôraz je kladený najmä na publikačné minimum štátnej správy a samosprávy a tzv. high-value datasety.

NKIVS konštatuje: „Dnes existuje veľký nesúlad medzi žiadanými otvorenými údajmi a publikovanými otvorenými údajmi. Na centrálnej platforme data.gov.sk prevládajú štatistické údaje, ktoré nemajú taký potenciál využitia v oblasti kontroly verejnej správy, alebo vytvárania nových IT riešení a aplikácií, ktoré by podporovali ekonomický rast a digitálnu ekonomiku.“ NKIVS, s. 29

NKIVS definuje 9 aktivít, ktoré by mali slúžiť k naplneniu cieľov.

Sú medzi nimi:

- Pokračovanie transponovania a implementácie smerníc Európskeho parlamentu súvisiacich s otvorenými údajmi,
- Zabezpečenie, aby verejná správa zverejňovala informácie štandardne vo forme otvorených údajov
- Zabezpečenie, aby otvorené údaje predstavovali aktívnu dátovú množinu údajov aj pre výkon verejnej správy,
- Publikovanie datasetov s vysokou hodnotou na opätovné použitie bezplatne a prostredníctvom API,
- Publikovanie verejnosťou žiadaných prioritných datasetov a zavedenie monitorovania úrovne využívania otvorených údajov.
- Poskytovanie metodickej, edukačnej, osvetovej, motivačnej a technickej podpory pre celú verejnú správu vrátane sprístupnenia technických riešení, ktoré môžu využívať na publikovanie dát.

³² <http://webcache.googleusercontent.com/search?q=cache:nBM3SUGnUdsJ:www.tretisektor.gov.sk/ktore-datasety-su-najzidanejsie/+&cd=2&hl=sk&ct=clnk&gl=sk&client=firefox-b-d>

- *Aplikovanie odporučených metodík, nástrojov, štandardov, princípov správy a rozvoja otvorených údajov meraných Európskou komisiou Indexom vyspelosti otvorených údajov (ODMI) v každom členskom štáte EU.*

NKIVS neobsahuje konkrétne merateľné ciele, čo by malo byť definované v Akčnom pláne k NKIVS, ktorý mal byť predložený na rokovanie vlády do marca 2020, čo sa dosiaľ nestalo.

5.5 Národný projekt Otvorené údaje 2.0

Národný projekt **Otvorené údaje 2.0 - Rozvoj centrálnych komponentov pre kvalitné zabezpečenie otvorených údajov** („OÚ 2.0“) nadväzuje na národný projekt Elektronické služby ÚV SR - eDemokracia a otvorená vláda prostredníctvom ktorého sa vytvoril portál otvorených dát data.gov.sk. Projekt „OÚ 2.0“ realizuje MIRRI a jeho zámerom je rozvoj a dobudovanie centrálnych technických komponentov a prostriedkov na trvalé sprístupnenie otvorených údajov verejnej správy SR.

V rámci projektu sa zabezpečí dostupnosť kľúčových otvorených datasetov (vďaka integrácii s platformou dátovej integrácie a kooperácií s dopytovými a rezortnými projektmi), dôveryhodnosť publikovaných údajov pre použitie digitálnej ekonomiky a archiváciu publikovaných údajov – posun k „trusted open data“. Zlepšia sa nástroje pre komunitu a používateľov a to vďaka lepšiemu vyhľadávaniu a nástrojov na zapájanie komunity. Publikovanie otvorených údajov bude zjednodušené. Zvýši sa kvalita publikovaných dát - cez využitie prepojených údajov (linked-data) a použitie Centrálného modelu údajov verejnej správy.

Rozšíria sa existujúce služby - vytvorenie SPARQL Endpointu, informovanie o spôsobe vyriešenia podnetu, definovanie metadát k novému datasetu alebo dátovému zdroju, podanie podnetu pre zlepšenie poskytovania otvorených dát, nadstavbové vyhľadávacie - sémantické.

Vytvoria sa nové služby - služba na publikáciu meta údajov do stránky, textové sémantické vyhľadávacie, SPARQL Endpoint – tj. Rozhranie pre fyzického človeka a stroj pre zadávanie dopytov na RDF databázu prostredníctvom jazyka SPARQL s podporou GeoSPARQL, Vyhodnocovanie kvality datasetu, monitoring a predikcia problematických datasetov, tvorba štatistík z prevádzky a vytvorenie repozitára zdrojových kódov a nahranie zdrojových kódov.

5.6 Výzva Manažment údajov inštitúcie verejnej správy³³

Účelom dopytovej výzvy je finančne podporiť aktivity, ktoré zabezpečia nastavenie systematického manažmentu údajov v jednotlivých organizáciách verejnej správy, ich **zdieľanie a publikovanie vo forme referenčných a otvorených údajov**, využívanie dát v rámci procesov inštitúcií, ako aj aktivity čistenia údajov pre dosiahnutie požadovanej kvality dát. Touto výzvou môžu byť podporené rôzne typy dopytovo – orientovaných projektov pre manažment údajov inštitúcií verejnej správy. Indikatívna výška finančných prostriedkov vyčlenených pre túto výzvu je 55 000 000

³³ <https://datalab.digital/dopytove-vyzvy/manazment-udajov-institucii-verejnejspravy/>

€, výzva bola vyhlásená 18.7.2019. V rámci výzvy sa realizujú alebo budú realizované tieto projekty³⁴:

- Manažment údajov RESS – Ministerstvo spravodlivosti
- Manažment údajov RPO
- Manažment údajov v prostredí regulácie sieťových odvetví
- Manažment údajov IS CITES
- Manažment údajov IS Register fluórovaných skleníkových plynov
- Konsolidácia, manažment a kvalita údajov VŠZP
- Manažment údajov Centrálnej archeologickej evidencie Slovenskej republiky
- Manažment údajov v oblasti hydromeliorácií
- Manažment údajov Dopravného úradu
- Manažment údajov v rezorte školstva
- Manažment údajov Ministerstva životného prostredia Slovenskej republiky
- Manažment údajov Úradu pre dohľad nad zdravotnou starostlivosťou - Centrálny register poisťencov, Register zdravotníckych pracovníkov, Register poskytovateľov zdravotnej starostlivosti, Register úmrtí
- Manažment údajov Zboru väzenskej a justičnej stráže
- Manažment údajov Ministerstva kultúry Slovenskej republiky
- Rozšírenie centrálného údajového repozitára ÚKSÚP - manažment údajov
- Manažment údajov pre oblasť vedy a výskumu
- Manažment údajov Úradu vlády Slovenskej republiky
- Manažment údajov Ministerstva financií Slovenskej republiky – Komplexný projekt pre manažment údajov
- Manažment údajov Informačného systému športu

Všetky projekty by mali obsahovať aj zverejňovanie otvorených dát alebo API. Projekty musia byť ukončené do konca programového obdobia (koniec roku 2023).

5.7 Akčný plán Iniciatívy pre otvorené vládnutie na roky 2022-2024

Akčný plán (AP) sa v čase písania tejto analýzy nachádza v medzirezortnom pripomienkovom konaní. Minulé akčné plány obsahovali viacero úloh spojených so zverejňovaním otvorených údajov, tento počet sa však časom zmenšoval.

Návrh AP obsahuje tieto 3 úlohy týkajúcu sa otvorených dát:

B. 3. v spolupráci so splnomocnencom vlády SR pre rozvoj občianskej spoločnosti participatívnym spôsobom vypracovať návrh metodiky mapovania dopadu otvorených dát v SR na vybrané oblasti,

do 31. decembra 2022,

³⁴ Okrem týchto projektov majú schválené projekty aj samosprávy: Trenčín, Snina, Svidník, Zlaté Moravce, Trnava, Zvolen, Liptovský Mikuláš, Žilinský samosprávny kraj

B. 4. v spolupráci so splnomocnencom vlády SR pre rozvoj občianskej spoločnosti zmapovať a vyhodnotiť kvalitatívne a kvantitatívne dáta vo vybraných oblastiach na základe vypracovanej metodiky a zmapovať a vyhodnotiť dopyt verejnosti po najžiadanejších otvorených dátach verejnej správy,

do 31. marca 2023,

B. 5. participatívnym spôsobom vypracovať správu o dopade otvorených dát v SR na vybrané oblasti a v spolupráci s relevantnými aktérmi pripraviť pravidelné meranie dopadu otvorených dát v SR,

do 30. septembra 2023

Za úlohy je zodpovedné Ministerstvo investícií, regionálnej rozvoja a informatizácie.

Cieľom záväzku je participatívnym spôsobom pripraviť metodiku na zhodnotenie dopadu otvorených dát na vybrané aspekty spoločnosti (napr. transparentnosť, vládnutie, ekonomiku, životné prostredie či na znevýhodnené komunity a.i), ako aj zhodnotiť relevantné otvorené dáta. Výstupom záväzku bude tiež správa o dopade otvorených dát a nastavenie pravidelného merania dopadu otvorených dát na vybrané aspekty spoločnosti, čím sa posilňujú kapacity aktívneho využívania dát, ako aj zavádza mechanizmy informovania o hodnote a prínose verejných prostriedkov pre občana.

5.8 Občianske združenie Alvaria

OZ Alvaria sa venuje projektom zameraným na otvorené dáta – od spracovávania a vizualizácie dát zo štátnych registrov, produktového vývoja a managementu, cez zapájanie komunity a dobrovoľníkov do vývoja IT riešení, až po tvorbu verejných politík v oblasti otvorených dát na národnej a samosprávnej úrovni. **OZ spája jednotlivcov a inštitúcie verejnej správy, akademické, neziskové a súkromné organizácie v Open Data Sieti³⁵**, aby spolu odhaľovali potenciál dát, identifikovali spoločenské výzvy a riešenia, premenili idey na konkrétne výsledky a spolu iniciovali pozitívnu spoločenskú zmenu. V septembri 2021 až marec 2022 realizovalo OZ projekt **Open Data Lab** – inkubátor pre študentov a pedagógov VŠ³⁶. Do projektu bolo celkovo zapojených viac ako 200 študentov a 13 pedagógov, zrealizovalo sa 17 workshopov. Väčšina workshopov bolo zameraných na vzdelávanie študentov a tvorbu študentských projektov nad otvorenými dátami. Projekt bol ukončený Demo Dňom, kde študenti odprezentovali svoje projekty expertom a zástupcom samospráv.

Z projektu vyplynulo, že vzdelávanie v oblasti vizualizácie a spracovania dát je na slovenských univerzitách žiadané, rovnako ako otvorené dáta a ich využitie pri výučbe a študentských projektoch. Výzvou ostáva tieto projekty ďalej monitorovať a šíriť informácie o nich, aby sa dostali k čo najviac samosprávam a orgánom verejnej a štátnej správy, ktoré následne získajú motiváciu k zverejňovaniu ďalších dát.

³⁵ <https://www.alvaria.sk/open-data-siet/>

³⁶ <https://www.alvaria.sk/otvorene-data-na-univerzitach-univerzity-chcu-pracovat-s-realnymi-datami-a-pripravovat-studentov-pre-prax/>

Okrem OZ Alvaria sa z občianskej spoločnosti témou otvorených údajov zaoberajú aj neziskové organizácie Open Data³⁷ a Slovenská spoločnosť pre otvorené informačné technológie (SOIT)³⁸. Open Data je skupina ľudí, ktorá chce presadiť otvorenú a modernú verejnú správu. SOIT je združenie ľudí, ktorých spája idea propagácie používania otvorených technológií v rôznych oblastiach života spoločnosti.

5.9 Zhrnutie

Za posledný rok **bola prijatá aktualizácia Národnej koncepcie informatizácie verejnej správy**, ktorá sa venuje aj otvoreným dátam. V oblasti otvorených údajov (strategická priorita Manažment údajov) je definovaný **princíp „čo nie je tajné je verejné“**, **dôraz je kladený najmä na publikačné minimum štátnej správy a samosprávy a tzv. high-value datasety**.

Bolo zverejnené publikačné minimum štátnej správy, OVM už začali na jeho základe na portáli zverejňovať datasety. Pomerne veľká časť datasetov je však naviazaná na schválenie zákona o údajoch a OVM zatiaľ nemajú povinnosť tieto datasety zverejňovať.

Dôležitý záväzok **Mapovanie vplyvu otvorených dát na vybrané aspekty spoločnosti** sa nachádza v novom, zatiaľ neschválenom Akčnom pláne Iniciatívy pre otvorené vládnutie na roky 2022-2024. Táto úloha by okrem zlepšenia pozície Slovenska v rebríčku Open Data Maturity mala pomôcť lepšie pochopiť dopady otvorených dát.

Projekt **OZ Alvaria Open Data Lab – inkubátor pre vysoké školy**, kde si študenti vytvárali nad dátami vlastné vizualizácie potvrdil záujem VŠ o prácu s otvorenými dátami.

Zverejňovanie datasetov a API je veľkým problémom, **zverejnených bolo len 22,8 % dopytovaných dát**, čo je veľmi málo.

Projekt Otvorené dáta 2.0 prinesie dostupnosť kľúčových otvorených datasetov, dôveryhodnosť publikovaných údajov pre použitie digitálnej ekonomiky a archiváciu publikovaných údajov – posun k „trusted open data“. Zlepšia sa nástroje pre komunitu a používateľov a to vďaka lepšiemu vyhľadávaniu a nástrojov na zapájanie komunity. Publikovanie otvorených údajov bude zjednodušené. Zvýši sa kvalita publikovaných dát - cez využitie prepojených údajov (linked-data) a použitie Centrálného modelu údajov verejnej správy.

³⁷ <https://opendata.sk/>

³⁸ <https://www.soit.sk/>

6 Záver

Centrálny portál otvorených dát data.gov.sk oslavuje už 10 rokov svojej existencie. Počas tejto doby sa zmenil len minimálne. Koncept a dizajn portálu už zastaral, navyše je zrejmé, že nedokázal podporiť vznik živej a rozsiahlej komunity okolo otvorených dát. Na túto realitu reaguje projekt OpenData 2.0, ktorý prinesie nový portál. Najväčšou výzvou však nebude vytvoriť samotný nový portál, ale poučiť sa z týchto chýb. Aby sa Slovensko posunulo ďalej v téme otvorených dát, potrebuje interaktívny a *otvorený* portál, ktorý bude pružnejšie reagovať na požiadavky používateľov a najmä umožní prepájať a podporovať spoluprácu producentov a používateľov dát.

7 Odporúčania

7.1 Opakujúce sa odporúčania z poslednej analýzy

1. **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.
2. **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 2021 pre Slovensko (kap. 3.3). Preto je užitočné realizovať projekt Open Data 2.0.

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

3. 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 zobrazit' všade aktuálny stav datasetov. Rozdiel môže byť spôsobený aj nejednoznačným zadaním, čo je dataset, distribúcia datasetu, formát datasetu, dátový zdroj, URI na dátový zdroj.
4. Je užitočné, že portál data.gov.sk zverejňuje datasety od každej organizácie zvlášť. Avšak bolo by prehľadnejšie, ak by portál data.gov.sk umožňoval zobrazit' zoznam ústredných orgánov štátnej správy³⁹ a k nim priradené podriadené organizácie⁴⁰. Napríklad Regionálne úrady verejného zdravotníctva priradené k Ministerstvu zdravotníctva SR. Tiež všetky územné samosprávy združit' do jedného zoznamu. Niektoré organizácie na portáli data.gov.sk majú nepresné alebo duplicitné názvy.
5. 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.
6. 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.
7. 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

³⁹https://www.slovensko.sk/sk/agendy/agenda/_ministerstva-a-ostatne-ustredne1

⁴⁰<https://www.slovensko.sk/sk/institucie>

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 odlišené od súborov na data.gov.sk.

8. Existuje zrozumiteľné hodnotenie otvorených dát hviezdčkami⁴¹. Kvalita poskytovaného datasetu je v tomto význame zadaná vo vyhláške ÚPVII č. 78/2020 Z. z. (Vyhláška o štandardoch pre informačné technológie verejnej správy)⁴², § 38 až § 40. Na zverejňovanie otvorených dát sa vzťahuje tiež Národná koncepcia informatizácie verejnej správy⁴³, Stratégia a akčný plán prístupnosti a používania otvorených údajov verejnej správy⁴⁴. Minimálna požadovaná úroveň pre otvorené dáta je úroveň 3 hviezdčiek. Portál data.gov.sk by mohol automaticky hodnotiť zverejnené datasety hviezdčkami 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.
9. Zaviesť prístupenie 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.
10. 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.
11. Portál data.gov.sk by mohol zobrazovať štatistiku využitia jednotlivých podaní⁴⁵ verejnosťou.
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 a vyhľadávať aplikácie aj na iných portáloch⁴⁶.
14. V rámci projektu Open Data 2.0 prejsť na CKAN API v najnovšej verzii 3.
15. 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.
16. 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.

⁴¹<https://5stardata.info/cs/>

⁴²<https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2020/78/>

⁴³<https://www.mirri.gov.sk/sekcie/informatizacia/narodna-koncepcia-informatizacie-verejnej-spravy-nikvs/index.html>

⁴⁴https://www.slovensko.sk/_img/CMS4/strategia_otvorene_udaje.pdf

⁴⁵<https://data.gov.sk/published-services>

⁴⁶ <https://utopia.sk/wiki/display/opendata/SK+a+CZ+opendata+aplikacie+a+projekty>

Odporúčania pre dátovú kanceláriu MIRRI

17. Po vzore ČR⁴⁷ monitorovať „časté zlé praktiky“ a zverejňovať krátke návody, popisujúce konkrétny nedostatok a kroky potrebné pre nápravu. Pozn.: MIRRI má vlastné návody⁴⁸, ktoré však nie sú natoľko prehľadné v porovnaní s ČR a neobsahujú komplexný zoznam najčastejších chýb.

Odporúčanie pre povinné osoby – poskytovateľov datasetov

18. Pred publikovaním datasetov odporúčame *preštudovanie Metodickéj príručky pre povinné osoby – ako zverejňovať otvorené dáta (2015)*⁴⁹, čím by sa dalo vyhnúť nedostatkom opísaným v tejto analýze.

7.2 Nové odporúčania

1. Zabezpečiť prioritné zverejnenie datasetov a API, ktoré vzišli z prieskumov dopytu verejnosti po najžiadanejších datasetoch a API.
2. Realizovať ďalší prieskum dopytu verejnosti po najžiadanejších datasetoch a API po splnení predchádzajúcej úlohy minimálne na 75 %.
3. Pravidelne vyhodnocovať úroveň zverejnenia najžiadanejších datasetov a API.
4. Pri realizácii prieskumu dopytu verejnosti po najžiadanejších datasetoch umožniť OVM a dátovým kurátorom kontaktovať respondentov, ktorí dobrovoľne na seba uvedú kontakt, v prípade nejasností.
5. Zverejniť návrhy na zverejnenie údajov na portáli otvorených dát v anonymizovanej podobe ako dataset v pravidelnej periodicite najmenej raz ročne.
6. Znižovať podiel datasetov bez uvedenia licencie (hodnota „iná“), datasetov s nedostatočne otvorenou licenciou (napr. CC-SA, CC-NC), okrem odôvodnených prípadov.
7. Sledovať využívanie (dopad) datasetov zverejnených na základe publikačného minima štátnej správy.
8. Vyhodnocovať dopyt verejnosti po datasetoch a API územnej samosprávy.

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

9. Publikovať manuál a/alebo videonávod ako napojiť existujúce portály otvorených dát automatizovane na centrálny portál a v prípade záujmu OVM k tomuto usporiadať workshop.
10. Zaviesť striktné požiadavky na zverejnené datasety – portál data.gov.sk by mal umožniť zverejnenie len tých datasetov, ktoré splnia aspoň 3 hviezdičky kvality. Odporúčame ponechať

⁴⁷<https://opendata.gov.cz/špatná-praxe:start>

⁴⁸ <https://wiki.vicpremier.gov.sk/pages/viewpage.action?pageId=82027683>

⁴⁹https://www.slovensko.sk/_img/CMS4/Navody/Metodicka_prirucka_pre_povinne_osoby.pdf



Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2022

OVM dostatočne dlhé prechodné obdobie, v ktorom by nevyhovujúce datasety boli nahradené vyhovujúcimi, aby sa tak žiadne údaje nestratili a neznižoval sa počet datasetov.

11. Umožniť automatickú konverziu formátov na portáli, t.j. nahraný súbor bude možné stiahnuť aj v inom formáte. Toto je dôležité pre pohodlie používateľov aj OVM. Niektorí používatelia nevedia pracovať s formátmi ako sú JSON, CSV, XML, pre nich je vhodné umožniť aj stiahnutie vo formáte XLSX/XLS/ODS.
12. Umožniť používateľovi tlačidlom a formulárom bez nutnosti elektronického podpisu nahlásiť chybu pri datasete. Oznamy by sa mali odosielať dátovým kurátorom.
13. Umožniť používateľovi podať návrh na zverejnenie údajov na portáli otvorených dát aj formou e-mailu alebo vo formulári bez nutnosti elektronického podpisu.
14. Zaviesť technický spôsob, ako rozlíšiť datasety s vysokou pridanou hodnotou na portáli od ostatných, napr. špeciálnym príznakom alebo tagom.
15. Umožniť používateľom filtrovať datasety aj podľa periódy aktualizácie.